

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 ; Search time 1380 Seconds  
(without alignments)  
1778.679 Million cell updates/sec

Title: US-09-475-704A-1

Perfect score: 60  
Sequence: 1 gacatcaagcagggccccc...tagacgcgtcttcaagacc 60

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481385 residues

Total number of hits satisfying chosen parameters: 1520254

Minimum DB seq length: C  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_sv.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_hgt\_hum.\*
- 31: em\_hgt\_inv.\*
- 32: em\_hgt\_other.\*
- 33: em\_hgt\_mus.\*
- 34: em\_hgt\_pli.\*
- 35: em\_hgt\_rod.\*
- 36: em\_hgt\_man.\*
- 37: em\_hgt\_vrt.\*
- 38: em\_sy.\*
- 39: em\_hgtq\_hum.\*
- 40: em\_hgtc\_mus.\*
- 41: em\_hgtgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	6	AX455885 Sequence
2	60	100.0	60	6	AX455937 Sequence
3	55.2	92.0	60	6	AX455886 Sequence
4	37.6	62.7	60	6	AX455930 Sequence
5	36	60.0	60	6	AX455938 Sequence
6	21.2	35.3	91	11	AF275608 AF275608 Bos taurus
7	20.4	34.0	60	6	AR169921 Sequence
8	20	33.3	81	6	AX480659 Sequence
9	19.8	33.0	97	1	AF223423 Pseudomon
10	19.6	32.7	84	9	HS91013 Sequence
11	19.4	32.3	78	6	AX480660 Sequence
12	19.2	32.0	51	6	AX158488 Sequence
13	19.2	32.0	97	6	AX088801 Sequence
14	19	31.7	53	6	I08503 Sequence 16
15	19	31.7	65	6	I08502 Sequence 17
16	18.8	31.3	30	6	AX026570 Sequence
17	18.8	31.3	64	6	AX453567 Sequence
18	18.8	31.3	74	5	AY227159 Aica toid
19	18.6	31.0	47	6	AX195033 Sequence
20	18.6	31.0	51	6	AX163045 Sequence
21	18.4	30.7	51	9	HSN4SPAT
22	18.4	30.7	62	6	AI0076 X07531 EBV B95-8 C
23	18.4	30.7	62	14	FBVCLGAG X07531 EBV B95-8 C
24	18.4	30.7	69	10	AY205871 Mas Muscu
25	18.4	30.7	74	6	AX098157 Sequence
26	18	30.0	80	14	AF040847 AF040847 Hepatitis
27	18	30.0	80	14	AF040849 Hepatitis
28	18	30.0	98	3	ASBP23A2 AF134278 Agrotis s
29	18	30.0	98	3	ASBP23B2 AF134280 Agrotis s
30	18	30.0	98	3	ASBP24A2 AF134282 Agrotis s
31	18	30.0	98	3	ASBP25A2 AF134284 Agrotis s
32	18	30.0	98	3	ASBP25B2 AF134286 Agrotis s
33	18	30.0	98	3	ASBP27A2 AF134288 Agrotis s
34	18	30.0	100	6	AR078713 Sequence
35	18	30.0	100	6	AR198693 Sequence
36	18	30.0	100	6	I25137 Sequence 33
37	17.8	29.7	40	6	AR142031 Sequence
38	17.8	29.7	40	6	I59898 Sequence 25
39	17.8	29.7	40	6	I68065 Sequence 1
40	17.8	29.7	40	6	I68756 Sequence 25
41	17.8	29.7	40	6	I55781 Sequence 25
42	17.8	29.7	47	6	AX194662 Sequence
43	17.8	29.7	51	9	S78515 type I kor
44	17.8	29.7	56	6	AX103647 Sequence
45	17.8	29.7	63	6	AR310621 Sequence

ALIGNMENTS

RESULT 1  
AX455885  
LOCUS AX455885 60 bp DNA linear PAT 06-JUL-2002  
DEFINITION Sequence 1 from Patent WO0204493.  
ACCESSION AX455885  
VERSION AX455885.1 GI:21714878  
KEYWORDS Human immunodeficiency virus  
SOURCE Human immunodeficiency virus  
ORGANISM Human immunodeficiency virus  
VIRUSES; Retroviral viruses; Retroviridae; Lentivirus; Primate  
IDENTIFICATION  
REFERENCE  
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.  
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,  
polypeptides and uses thereof

```
JOURNAL Patent: WO 0204493-A 1 17-JAN-2002;
FEATURES source
Location/Qualifiers
1..60
/organism="Human immunodeficiency virus"
/mol_type="genomic DNA"
/db_xref="taxon:12721"
BASE COUNT 14 a 22 c 15 g 9 t
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGCCCCCAAGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAAGACC 60
Db 1 GACATCAAGCAGGCCCCCAAGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAAGACC 60

RESULT 2
AX455937
LOCUS AX455937
DEFINITION Sequence 53 from Patent WO0204493.
ACCESSION AX455937
VERSION AX455937.1 GI:21714921
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL Patent: WO 0204493-A 53 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES source
1..60
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/db_note="HIV Type C Gag Major Homology Region Optimized"
BASE COUNT 14 a 22 c 15 g 9 t
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGCCCCCAAGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAAGACC 60
Db 1 GACATCAAGCAGGCCCCCAAGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAAGACC 60

RESULT 3
AX455886
LOCUS AX455886
DEFINITION Sequence 2 from Patent WO0204493.
ACCESSION AX455886
VERSION AX455886.1 GI:21714879
KEYWORDS
SOURCE Human immunodeficiency virus
ORGANISM Human immunodeficiency virus
Viruses; Retroviral viruses; Retroviridae; Lentivirus; Primate
lentivirus group.
REFERENCE
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL Patent: WO 0204493-A 2 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES source
1..60
/organism="Human immunodeficiency virus"
/mol_type="genomic DNA"
/db_xref="taxon:12721"
/db_note="HIV Type C Gag Major Homology Region Wild Type"
BASE COUNT 24 a 13 c 11 g 12 t
ORIGIN

JOURNAL Patent: WO 0204493-A 1 17-JAN-2002;
FEATURES source
Location/Qualifiers
1..60
/organism="Human immunodeficiency virus"
/mol_type="genomic DNA"
/db_xref="taxon:12721"
BASE COUNT 12 a 24 c 15 g 9 t
ORIGIN
Query Match 92.0%; Score 55.2; DB 6; Length 60;
Best Local Similarity 95.0%; Pred. No. 7e-06;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGCCCCCAAGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAAGACC 60
Db 1 GACATCAAGCAGGCCCCCAAGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAAGACC 60

RESULT 4
AX455910
LOCUS AX455910
DEFINITION Sequence 26 from Patent WO0204493.
ACCESSION AX455910
VERSION AX455910.1 GI:21714898
KEYWORDS
SOURCE Human immunodeficiency virus
ORGANISM Human immunodeficiency virus
Viruses; Retroviral viruses; Retroviridae; Lentivirus; Primate
lentivirus group.
REFERENCE
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL Patent: WO 0204493-A 26 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES source
1..60
/organism="Human immunodeficiency virus"
/mol_type="genomic DNA"
/db_xref="taxon:12721"
BASE COUNT 24 a 13 c 11 g 12 t
ORIGIN
Query Match 62.7%; Score 37.6; DB 6; Length 60;
Best Local Similarity 76.7%; Pred. No. 0.97;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGCCCCCAAGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAAGACC 60
Db 1 GACATCAAGCAGGCCCCCAAGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAAGACC 60

RESULT 5
AX455938
LOCUS AX455938
DEFINITION Sequence 54 from Patent WO0204493.
ACCESSION AX455938
VERSION AX455938.1 GI:21714922
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL Patent: WO 0204493-A 54 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES source
1..60
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/db_note="HIV Type C Gag Major Homology Region Wild Type"
BASE COUNT 24 a 13 c 11 g 12 t
ORIGIN
```

Query Match	34.0%	Score 20.4;	DB 6;	Length 60;
Best Local Similarity	55.2%;	Pred. No. 1e+05;		
Matches	30;	Conservative	0;	Mismatches 16; Indels 0; Gaps 0;
QY	13	GGCCCCAAGGAGCCCTTCGCGCACTACGGAGCCGCTCTTCAAGA	58	
DB	13	GGUCCTTAGAACCTTTTAGAGACATGTTGATAGGTTTATAGA	58	
RESULT 8				
AX480659		81 bp	DNA	linear
LOCUS				
DEFINITION		Sequence 47 from Patent WO0248189.		
ACCESSION		AX480659		
VERSION		AX480659.1	GI:22217408	
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS		Etzerodt, M., Holtet, I. L., Graversen, N. J., and Th. Gersen, H. C.		
TITLE		Combinatorial libraries of proteins having the scaffold structure		
JOURNAL		Patent: WO 0248189-A 47 20-JUN-2002;		
		Borean Pharma A/S (DK)		
FEATURES				
SOURCE		1..81		
		/organism="synthetic construct"		
		/mol_type="genomic DNA"		
		/db_xref="taxon:32530"		
		/note="oligonucleotide"		
BASE COUNT		12 a	20 c	22 g
ORIGIN				18 others
Query Match	33.3%	Score 20;	DB 6;	Length 81;
Best Local Similarity	44.2%;	Pred. No. 1.3e+05;		
Matches	23;	Conservative	6;	Mismatches 23; Indels 0; Gaps 0;
QY	1	GACATCAAGCAGGCGCCCAAGGAGCGCTTCGCGCACTACGTCACCGCTCT	52	
DB	2	GAGATCTGGCTGGCCCTCAGCAGCNNSNNNSNNNSNNSGAGCGACCT	53	
RESULT 9				
AF223423		97 bp	DNA	linear
LOCUS				
DEFINITION		Pseudomonas aeruginosa clone 570 biofilm-dependent; regulatory		
ACCESSION		AF223423		
VERSION		AF223423.1	GI:7340077	
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS		Weyers, J. B., van Zyl, W. H., and Brozel, V. S.		
TITLE		Pseudomonas aeruginosa contains a novel and diverse group of		
		biofilm-dependent regulatory elements		
JOURNAL		Unpublished		
REFERENCE				
AUTHORS		Weyers, J. B., and Brozel, V. S.		
TITLE		Direct Submission		
JOURNAL		Submitted (11-JAN-2000) Microbiology and Plant Pathology,		
		University of Pretoria, Lunnon Road, Pretoria, Gauteng 0002, South		
FEATURES				
SOURCE		1..97		
		/organism="Pseudomonas aeruginosa"		
		/mol_type="genomic DNA"		
		/strain="PAOI"		
		/db_xref="taxon:287"		
		/clone="570"		

```

misc_feature      1..97
/note="biofilm-dependent regulatory element; up-regulated
when grown as a biofilm on a solid surface"
BASE COUNT      16 a      35 c      11 t
ORIGIN
Query Match      33.0%; Score 19.8; DB 1; Length 97;
Best Local Similarity 63.8%; Pred. No. 1.4e+05;
Matches 30; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 9 GCAGGGCCCCAAGAGCGCTTCCCGGACTACGTGGACCGCTTCTCA 55
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25 GCGCGGATCGAGGTGCGGTTCGACGACGCGGCGGCGCTGTCTCA 71

RESULT 10
HS091013/c
LOCUS      HS091013      84 bp      mRNA      linear      PRI 15-JUL-1997
DEFINITION Homo sapiens clone FEL399 T-cell receptor delta chain (TCRVDL3);
            mRNA, partial cds.
ACCESSION  U91013
VERSION     U91013.1 GI:2239433
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1. (bases 1 to 84)
AUTHORS    Holtmeier,W., Witthoft,T., Hennemann,A., Winter,H.S. and
            Kagnoff,M.F.
TITLE      The TCR-delta repertoire in human intestine undergoes
            characteristic changes during fetal to adult development
JOURNAL    J. Immunol. 158 (12). 5632-5641 (1997)
MEDLINE    97334214
PUBMED     9190911
REFERENCE  2 (bases 1 to 84)
            Holtmeier,W., Witthoft,T., Hennemann,A., Harland,S.W. and
            Kagnoff,M.F.
            Direct Submission
TITLE      Submitted (27-FEB-1997) Department of Medicine, University of
            Frankfurt, Theodor-Stern Kai #7, Frankfurt 60590, Germany
JOURNAL
FEATURES   Location/Qualifiers
            source
            1..84
            /organism="Homo sapiens"
            /mol_type="RNA"
            /db_xref="taxon:9606"
            /chromosome="14"
            /clone="FEL399"
            /tissue_type="small intestine"
            /dev_stage="fetus; 20 weeks gestation"
            <1..>84
            /gene="TCRVDL3"
            <1..>84
            /gene="TCRVDL3"
            /note="rearranged; contains CDR3 domain, 16 amino acids
            (calculation according to: Rock,K., J. Exp.Med.,
            179:323-328,1994)"
            /codon_start=1
            /product="T-cell receptor delta chain"
            /protein_id="AAC31575.1"
            /db_xref="GI:2239434"
            /translation="LEDSAKYFCALGEPTTWGIGSWTRQNF"
            <1..>84
            V_region
            /gene="TCRVDL3"
            20 a      18 c      25 g      21 t
ORIGIN
Query Match      32.7%; Score 19.6; DB 9; Length 84;
Best Local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 2 ACATCAAGCAGGGCCCCAAGAGCGCTTCCCGGACTACGTGG 43
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB #2 ACATCTGTGGGIGTCCAGAGCGCTATCCCGCACGTAAGTCCG 41
RESULT 11
AX180660
LOCUS      AX180660      78 bp      DNA      linear      PAT 12-AUG-2002
DEFINITION Sequence 48 from Patent WO0248185.
ACCESSION  AX480660
VERSION     AX480660.1 GI:22217409
KEYWORDS   synthetic construct
            synthetic construct
            artificial sequences.
ORGANISM   1
SOURCE     Elzerodt,M., Hollet,T.L., Graversen,N.J. and th Gersen,H.C.
            Combinatorial libraries of proteins having the scaffold structure
            of c-type lectin-like domains
            Patent: WO 0248189-A 48 20-JUN-2002;
            Borean Pharma A/S (DK)
            Location/Qualifiers
            source
            1..78
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
            /note="oligonucleotide"
BASE COUNT      12 a      20 c      22 g      9 t      15 others
ORIGIN
Query Match      32.3%; Score 19.4; DB 6; Length 78;
Best Local Similarity 46.5%; Pred. No. 1.9e+05;
Matches 23; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGGCCCCAAGAGCGCTTCCCGGACTACGTGGACCGCT 49
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 GAGATCTGGCTGGGCTCAACGACNNNNNNNNNNNSGAGGCGACCT 50

RESULT 12
AX158488/c
LOCUS      AX158488      51 bp      DNA      linear      PAT 22-JUN-2001
DEFINITION Sequence 1816 from Patent WO0140521.
ACCESSION  AX158488
VERSION     AX158488.1 GI:14539619
KEYWORDS   Homo sapiens (human);
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1
            Shinkens,R.A. and Leach,M.
            Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
            Patent: WO 0140521-A 1816 07-JUN-2001;
            Curagen Corporation (US)
            Location/Qualifiers
            source
            1..51
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            misc_feature
            26
            /note="2 of 2 allelic variants (1815 is other entry)"
            Accession: number cg35066497"
BASE COUNT      3 a      13 c      17 g      18 t
ORIGIN
Query Match      32.0%; Score 19.2; DB 6; Length 51;
Best Local Similarity 67.5%; Pred. No. 2.3e+05;
Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 17 CCAAGGAGCGCTTCCCGGACTACGTGGACCGCTTCTTCAA 56
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 47 CCAAGGACGCCATCAGCGACCGACCGAGGACGACATCAAGAA 8

```



```

Unclassified.
1 (bases 1 to 65)
AUTHORS Fiddes J.C. and Abraham J.A.
TITLE RECOMBINANT FIBROBLAST GROWTH FACTORS
JOURNAL Patent: WO 8701728-A 16-26-MAR-1987;
FEATURES Location/Qualifiers
source 1..65
/organism="unknown"
11 a 23 c 19 g 12 t
BASE COUNT
ORIGIN
Query Match: 31.7%; Score 19; DB: 6; Length 65;
Best Local Similarity 65.1%; Pred. No. 2.6e-05;
Matches 28; Conservative 0; Mismatches 15; Indels
QY 11 AGGGCCCCAAGGAGCGCTTCGGCGAGTACGTGGACCGCTTCTT 53
1111111111 11 11 11 1111111111
DB 20 AGGACCCCAAGCGGCTGTACTGCAAGAACGGGGGCTTCCTCT 62
1111111111 11 11 11 1111111111
Search completed: September 22, 2003, 14:56:51
Cpu time : 1.386 secs

```



```
Db 61 CTGCGCCCGGCGCAAGAAGCACTACATGCTGAAGCACTGGTGTGGGCCAGCCGTGAG 120
QY 121 CTGGAGGGTTTGGCCCTGAACCCCGCCCTGCTTGGAGACCGCCGAGGGTGTGAAGCAGATC 180
Db 121 CTGGAGGGTTTGGCCCTGAACCCCGCCCTGCTTGGAGACCGCCGAGGGTGTGAAGCAGATC 180
QY 181 ATGAACAGCTGTGAGCCGCGCTGTAGACCGGCACAGAGCACTGCGGAGCTGTACAA 240
Db 181 ATGAACAGCTGTGAGCCGCGCTGTAGACCGGCACAGAGCACTGCGGAGCTGTACAA 240
QY 241 ACGTGGCAACCTGTACTGCGGTGACGCGGCACTGAGAGTTCGGGACACCAAGAGGCC 300
Db 241 ACGTGGCAACCTGTACTGCGGTGACGCGGCACTGAGAGTTCGGGACACCAAGAGGCC 300
QY 301 CTGGCAACAGCTGAGGAGGAGGAGCAAGTCCGAGGCAAGACCCAGGAGGAGG 360
Db 301 CTGGCAACAGCTGAGGAGGAGGAGCAAGTCCGAGGCAAGACCCAGGAGGAGG 360
QY 361 GCGGAGGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTGCGAGGGCCAGATGGTG 420
Db 361 GCGGAGGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTGCGAGGGCCAGATGGTG 420
QY 421 CACGAGGCATCAGCCCGCCGACCTGACGAGGCTGGGTGAAGGTGATCAGGAGAGGCC 480
Db 421 CACGAGGCATCAGCCCGCCGACCTGACGAGGCTGGGTGAAGGTGATCAGGAGAGGCC 480
QY 481 TTGAGCCCGGAGGTGATCCGCACTGTTACCGGCCCTGAGCGAGGCGCCAGCCGAGAG 540
Db 481 TTGAGCCCGGAGGTGATCCGCACTGTTACCGGCCCTGAGCGAGGCGCCAGCCGAGAG 540
QY 541 CTGAACACAGATGTTGAACACCTGTTGGCGGCGCCAGGCGGCGCATGAGATGTGAAGAC 600
Db 541 CTGAACACAGATGTTGAACACCTGTTGGCGGCGCCAGGCGGCGCATGAGATGTGAAGAC 600
QY 601 ACCATCAAGAGAGGAGCGCCGAGTGGAGCGCTGACCGCGTGCAGCGCCGCGCCGCTG 660
Db 601 ACCATCAAGAGAGGAGCGCCGAGTGGAGCGCTGACCGCGTGCAGCGCCGCGCCGCTG 660
QY 661 GCGCCGCGCAGATGCGGACCGCCGCGAGCATGCGGCGGCGCCAGCATGCGGCGGCGCC 720
Db 661 GCGCCGCGCAGATGCGGACCGCCGCGAGCATGCGGCGGCGCCAGCATGCGGCGGCGCC 720
QY 721 CAGGAGCAGATCGCCTGGATGACCCAGCACCCCGCGAGCATGCGGCGGCGCCAGCATG 780
Db 721 CAGGAGCAGATCGCCTGGATGACCCAGCACCCCGCGAGCATGCGGCGGCGCCAGCATG 780
QY 781 CGGTGATCATCTGCGCTGAAACAGATGCTGCGGATGACAGCGCCGCTGAGATCTGCTG 840
Db 781 CGGTGATCATCTGCGCTGAAACAGATGCTGCGGATGACAGCGCCGCTGAGATCTGCTG 840
QY 841 GACATCGCGAGGCGCCGAGAGGCGCTTCGCGAGTACGTGGACCGCTTCTTCAAGACC 900
Db 841 GACATCGCGAGGCGCCGAGAGGCGCTTCGCGAGTACGTGGACCGCTTCTTCAAGACC 900
QY 901 CTGCGCGCGAGCAGCCGCGAGGAGCGTGAAGAACTGATGACGAGAGCGCTGCTGCTG 960
Db 901 CTGCGCGCGAGCAGCCGCGAGGAGCGTGAAGAACTGATGACGAGAGCGCTGCTGCTG 960
QY 961 CAGAAAGCCAAACCCGACTGCAAGAACATCTCTGCGGCTCTGCGGCGCGGCGCCAGCTG 1020
Db 961 CAGAAAGCCAAACCCGACTGCAAGAACATCTCTGCGGCTCTGCGGCGCGGCGCCAGCTG 1020
QY 1021 GAGGAGATGATGACCGCTTGGCAGGCGCTGGGCGCGCGCGCCAGCAAGGCCGCGTGTG 1080
Db 1021 GAGGAGATGATGACCGCTTGGCAGGCGCTGGGCGCGCGCGCCAGCAAGGCCGCGTGTG 1080
QY 1081 GCGAGGCGATGAGCCAGGCGCAACGCTGACATCATCATCATCATCATCATCATCATCAT 1140
Db 1081 GCGAGGCGATGAGCCAGGCGCAACGCTGACATCATCATCATCATCATCATCATCATCAT 1140
QY 1141 GCGCCGCGGCGCAACCTGAAGTCTTCAACTGCGCAAGGAGGCGGCACATGCGCAAGAAC 1200
Db 1141 GCGCCGCGGCGCAACCTGAAGTCTTCAACTGCGCAAGGAGGCGGCACATGCGCAAGAAC 1200
```

```
QY 1201 TGCGCGCCCGGCGCAAGAAGGCTGCTGAAAGTGGCGCAAGAGGGCCACCAGATGAAG 1260
Db 1201 TGCGCGCCCGGCGCAAGAAGGCTGCTGAAAGTGGCGCAAGAGGGCCACCAGATGAAG 1260
QY 1261 GACTGCAACCGAGCGCCAGGCGCAACTTCTTGGCAAGATCTTGGCTCAGCCCAAGGCGCG 1320
Db 1261 GACTGCAACCGAGCGCCAGGCGCAACTTCTTGGCTCAGCACTTGGCTCAGCCCAAGGCGCG 1320
QY 1321 CCGCGCAACTTCTTGCAGAACCGGAGGAGCGCCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1321 CCGCGCAACTTCTTGCAGAACCGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
QY 1381 CCGCGCGAGAGCTTCCGCTTGGAGGAGCACCCCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 1381 CCGCGCGAGAGCTTCCGCTTGGAGGAGCACCCCGCGCGCGCGCGCGCGCGCGCGCG 1440
QY 1441 CCGAGCGCTTACCGCGAGCGCGCTGACCGCGCTGCGAGCGCTTTCGCGCAGCGCGCGCG 1500
Db 1441 CCGAGCGCTTACCGCGAGCGCGCTGACCGCGCTGCGAGCGCTTTCGCGCAGCGCGCGCG 1500
QY 1501 AGCCAGTAA 1509
Db 1501 AGCCAGTAA 1509
```

RESULT 2  
US-09-967-464-68  
; Sequence 68, Application US/09967464  
; Publication No. US20030138453A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Hagan, Derek  
; APPLICANT: Otten, Gillis  
; APPLICANT: Donnelly, John J.  
; APPLICANT: Polo, John M.  
; APPLICANT: Barnett, Susan  
; APPLICANT: Singh, Mamohan  
; APPLICANT: Ulmer, Jeffrey  
; APPLICANT: Dubensky, Jr., Thomas W.  
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS  
; FILE REFERENCE: P16269.004  
; CURRENT APPLICATION NUMBER: US/09/967,464  
; CURRENT FILING DATE: 2002-04-11  
; PRIOR APPLICATION NUMBER: 60/236,105  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: 60/315,905  
; PRIOR FILING DATE: 2001-08-30  
; NUMBER OF SEQ ID NOS: 68  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 68  
; LENGTH: 1509  
; TYPE: DNA  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-967-464-68

Query Match 99.78; Score 1504.2; DB 12; Length 1509;  
Resid. Local Similarity 99.28; Pred. No. 8.5e-299;  
Matches 1497; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGCGCGCGCGCGCGCGCATCTCGCGCGCGGAGAAAGCTGGACAAAGTGGGAGAAAGATCCGC 60
Db 1 ATGGCGCGCGCGCGCGCGCATCTCGCGCGCGGAGAAAGCTGGACAAAGTGGGAGAAAGATCCGC 60
QY 61 CTGGCGCGCGCGCGCGCGCGCTGCTGAGAGACCTGCTGAGAGACCTGCTGCTGGCGCGCGGAG 120
Db 61 CTGGCGCGCGCGCGCGCGCGCTGCTGAGAGACCTGCTGAGAGACCTGCTGCTGGCGCGCGGAG 120
QY 121 CTGGAGGCTTCCGCTGAACCGCGCGCTGCTGGAGACCGCGCGGAGCGCTGCAAGCAGATC 180
Db 121 CTGGAGGCTTCCGCTGAACCGCGCGCTGCTGGAGACCGCGCGGAGCGCTGCAAGCAGATC 180
QY 181 ATGAAGCAGTGTGAGCGCGCGCGCTGACAGCGGAGCGCGCGAGGCTGGCGAGCGCTGTACAAC 240
Db 181 ATGAAGCAGTGTGAGCGCGCGCGCTGACAGCGGAGCGCGCGAGGCTGGCGAGCGCTGTACAAC 240
```









Db 847 CTGGACATCAAGCAGGCGCCCAAGAGCCCTTCGGGACTAGCTGGACGCGCTTCTTCAAG 936  
Qy 898 ACCCTGGCGCCGAGCAGGCGCCCAAGAGCCCTTCGGGACTAGCTGGACGCGCTTCTTCAAG 937  
Db 907 ACCCTGGCGCCGAGCAGGCGCCCAAGAGCCCTTCGGGACTAGCTGGACGCGCTTCTTCAAG 938  
Qy 958 GTGCAGAACCCCAACCCGAGCTGCAAGACCATCTCTGGCGCTCTCTGGCGCCCGGCGCCAC 1017  
Db 967 GTGCAGAACCCCAACCCGAGCTGCAAGACCATCTCTGGCGCTCTCTGGCGCCCGGCGCCAC 1026  
Qy 1018 CTGAGGAGATGATGACCGCGCTGCGAGGCGCTGGCGGCGCCCGGCGCCCAAGAGCGCGGATG 1077  
Db 1027 CTGAGGAGATGATGACCGCGCTGCGAGGCGCTGGCGGCGCCCGGCGCCCAAGAGCGCGGATG 1086  
Qy 1078 CTGCGCCGAGGATGAGCGGCGCCCAAGACCGCTGACATCATGATGCAAGAGCAATTC 1137  
Db 1087 CTGCGCCGAGGATGAGCGGCGCCCAAGACCGCTGACATCATGATGCAAGAGCAATTC 1143  
Qy 1138 AAGGCGCCCGGCGCCCAAGCTCAAGTGTCTCAACTGCGGCAAGGAGGCGCCACATCGCGCAAG 1197  
Db 1144 AAGGCGCCCGGCGCCCAAGCTCAAGTGTCTCAACTGCGGCAAGGAGGCGCCACATCGCGCC 1203  
Qy 1198 AACTGCGCGCGCCCGGCGCCCAAGAGGCTGTGTGAATGTGGCGCAAGAGGCGCCACATGATG 1257  
Db 1204 AACTGCGCGCGCCCGGCGCCCAAGAGGCTGTGTGAATGTGGCGCAAGAGGCGCCACATGATG 1263  
Qy 1258 AAGGACTGCAACCGGCGCGGCGCCCAACTTCTGGCAAGATCTGGCGCAAGAGGCGCCACAGGCG 1317  
Db 1264 AAGGACTGCAACCGGCGCGGCGCCCAACTTCTGGCAAGATCTGGCGCAAGAGGCGCCACAGGCG 1323  
Qy 1318 CGCCCGCGCAACTTCTGTGCAAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377  
Db 1324 CGCCCGCGCAACTTCTGTGCAAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1365  
Qy 1378 CGCCCGCGGAGAGCTTCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1437  
Db 1366 CGCCCGCGGAGAGCTTCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1425  
Qy 1438 GACCGGAGGCGCTACCGCGGAGCG 1497  
Db 1426 G-----ACCGGAGAGCGCTGACCGAGCGCTGAGAGGCGCTTCTGGCAAGGAGCGCG 1475  
Qy 1498 CTGAGCCAGTAA 1509  
Db 1474 CTGAGCCAGAAA 1485

## RESULT 7

US-10-190-435-35  
; Sequence 35, Application US/10190435  
; Publication No. US2003014324A1  
; GENERAL INFORMATION:  
; APPLICANT: ZUR MEDEDE, Jac  
; APPLICANT: BARNETT, Susan W.  
; APPLICANT: LIAN, Ying  
; APPLICANT: ENGELBRECHT, Susan  
; APPLICANT: VAN RENSBURG, Estrelita J.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
; FILE REFERENCE: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: P18133.003 / 2302-18133  
; CURRENT APPLICATION NUMBER: US/10/190.435  
; CURRENT FILING DATE: 2002-12-30  
; NUMBER OF SEQ ID NOS: 319  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 35  
; LENGTH: 4483  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: gp160mod.TV1.dV1V2-gagmod.BW965

Query Match 84.5%; Score 1275.2; DB 12; Length 4483;  
Sect Local Similarity 92.5%; Pred. No. 5.4e-252;  
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;  
Qy 1 ATGGGCG 3064  
Db 3005 ATGGGCG 3064  
Qy 61 CTGCG 120  
Db 3065 CTGCG 3124  
Qy 121 CTGAGGCGCTTCTGGCGCTGAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180  
Db 3125 CTGAGGCGCTTCTGGCGCTGAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3184  
Qy 181 ATGAACAGCTGTGAGCG 240  
Db 3185 ATCGGCGAGCTGACCG 3244  
Qy 241 ACCGTGCG 300  
Db 3245 ACCGTGCG 3304  
Qy 301 CTGCAACAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
Db 3305 CTGCAACAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3364  
Qy 361 GCGGAC---GGCAAGGTGAGCG 417  
Db 3365 GCGGACAAAGGCGAAGGTGAGCG 3424  
Qy 418 GTGCAACAGGCGCATGAGCG 477  
Db 3425 GTGCAACAGGCGCATGAGCG 3484  
Qy 478 GCGTTGAGCG 537  
Db 3485 GCGTTGAGCG 3544  
Qy 538 GAGCTGAACAGATGTGACACCGGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597  
Db 3545 GAGCTGAACAGATGTGACACCGGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3604  
Qy 598 GACACCATCAACGAGGAGGCG 657  
Db 3605 GACACCATCAACGAGGAGGCG 3664  
Qy 658 GTGGCG 717  
Db 1665 ATGCG 3724  
Qy 718 CTGAGGAGGAGATCGCGCTGATGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 777  
Db 3725 CTGAGGAGGAGATCGCGCTGATGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3784  
Qy 778 AAGCGGTGGATCATCTCTGGCGCTGAAACAGATCTGCGGATGTACAGCGCGCGCGCGCGCGCG 837  
Db 3785 AAGCGGTGGATCATCTCTGGCGCTGAAACAGATCTGCGGATGTACAGCGCGCGCGCGCGCGCG 3844  
Qy 838 CTGAGATCTCG 897  
Db 3845 CTGAGATCTAAGCG 3904  
Qy 898 ACCGTGCG 957  
Db 3905 ACCGTGCG 3964  
Qy 958 GTGCAAGACCGCAACCG 1017  
Db 3965 GTGCAAGACCGCAACCG 4024  
Qy 1018 CTGAGGAGATGATGACCG 1077







QY 1378 CCCCCCGGAGAGCTTCGGTTCAGGAGACCAACCCCGCCGCCCAAGAGGAGCCCAAG 1437  
|||||  
Db 4496 CCCCCCGGAGAGCTTCGGTTCAGGAGACCAACCCCGCCGCCCAAGAGGAGCCCAAG 4555  
|||||  
QY 1438 CACCGGAGCCCTACCGGAGCCCTGACGCCCTGCGGAGCCCTGTTTCGGGAGGGCCCC 1497  
|||||  
Db 4556 G-----ACCGGAGACCCCTGACCAGCCTGAAGAGCCCTGTTTCGGGCAAGGACCCC 4603  
|||||  
QY 1498 CTGAGCCAGTAA 1509  
|||||  
Db 4604 CTGAGCCCAATA 4615  
|||||  
RESULT 10  
US-10-190-435-38  
; Sequence 38, Application US/10190435  
; Publication No. US20030143248A1  
; GENERAL INFORMATION:  
; APPLICANT: ZUR MEDEDE, Jan  
; APPLICANT: BARNETT, Susan W.  
; APPLICANT: LIAN, Ying  
; APPLICANT: ENGELBRECHT, Susan  
; APPLICANT: VAN RENSBERG, Estrelita J.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
; FILE REFERENCE: P2-813.003 / 2702-18133  
; CURRENT APPLICATION NUMBER: US/10190.435  
; CURRENT FILING DATE: 2002-12-30  
; NUMBER OF SEQ ID NOS: 319  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 38  
; LENGTH: 4702  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: gp160mod.TVI-gagmod.BW955  
US-10-190-435-38  
Query Match 84.5%; Score 1275.2; DB 12; Length 4702;  
Best Local Similarity 92.5%; Pred. No. 5.4e-252;  
Matches 1398; Conservative 0; Mismatches 75; Indels 36; Gaps 4;  
QY 1 ATGGGCGCCCGCCAGCATCTCTCGGGGAGAGAGTGGACACACTGGGAGAGATCGGC 60  
|||||  
Db 3224 ATGGGCGCCCGCCAGCATCTCTCGGGGAGAGAGTGGACACACTGGGAGAGATCGGC 3283  
|||||  
QY 61 CTGGCGCCCGCGGCAAGCACTACATGCTGAAGCACTTGTGTGGCGCAGCGGAG 120  
|||||  
Db 3284 CTGGCGCCCGCGGCAAGCAAGTGTATGATGAAGCACTTGTGTGGCGCAGCGGAG 3343  
|||||  
QY 121 CTGGAGGCTTCGCGCTGAACCCCGGCTGTGGAGACCGCGGAGGCTGCAAGCAGATC 190  
|||||  
Db 3344 CTGGAGAGTTCGCGCTGAACCCCGGCTGTGGAGACCGCGGAGGCTGCAAGCAGATC 3403  
|||||  
QY 181 ATGAAGCAGCTGACGCCGCCCTGCAGACCGGCAACGAGGAGTGGCCAGCTGTACAAC 240  
|||||  
Db 3404 ATCGGCCAGCTGCAACCGCGCTGCAGACCGGCAAGGAGGAGTGAAGAGCTTTCAC 3453  
|||||  
QY 241 ACCGTGGCCACCTGTACTCGTGCAGCGCGGATCGAGGTCCCGGACCAAGAGGCC 320  
|||||  
Db 3464 ACCGTGGCCACCTGTACTCGTGCAGAGAGATCGAGGTCCCGGACCAAGAGGCC 3523  
|||||  
QY 301 CTGGACAGATCGAGGAGGAGCAAAAGTCCCGCAGGAGAGACCCAGCGGCGCAAGGAG 350  
|||||  
Db 3524 CTGGACAGATCGAGGAGGAGCAAAAGTCCCGCAGGAGAGTCCAGCGGCGCAAGGAG 3583  
|||||  
QY 361 GCGAC-----GCGAAGGTGAGCAAACTACCCCTCTGTCAGAACTCTGAGGGGCGAGA 417  
|||||  
Db 3584 CCGCAGACAGGCAAGGTGAGCAAACTACCCCTCTGTCAGAACTCTGAGGGGCGAGA 4643  
|||||  
QY 418 GTGACACAGGCCATCAGCCCCCGCCACCTGTGAACGCCCTGGGTGAAGGTGATCGAGAGAG 477  
|||||

Db 3644 GTGACACAGGCCATCAGCCCCCGCCACCTGAACGCTGGGTGAAGCTGATCGAGGAGAAG 3703  
|||||  
QY 478 GCCTTCAGCCCCAGGTGATCCCAAGTTCACCGCCTGAGCGAGSGCGCCACCCCGCAG 537  
|||||  
Db 3704 GCCTTCAGCCCCAGGTGATCCCAAGTTCACCGCCTGAGCGAGSGCGCCACCCCGCAG 3763  
|||||  
QY 538 GACCTGAACACAGATGTTGAACACCGTGGGGGGCCACCAAGCGCCCATGAGATGCTGAAG 597  
|||||  
Db 3764 GACCTGAACACAGATGTTGAACACCGTGGGGGGCCACCAAGCGCCCATGAGATGCTGAAG 3823  
|||||  
QY 598 GACACCATCAACAGAGAGGCGCGGAGTGGGAGCGCTGCACCCCTGAGCGCGGCCGCC 657  
|||||  
Db 3824 GACACCATCAACAGAGAGGCGCGGAGTGGGAGCGCTGCACCCCTGAGCGCGGCCGCC 3883  
|||||  
QY 658 GTGGCCCCCGGCAGATGCGGACCCCGCGGAGGAGATCGCGCGCGCAACAGCACCC 717  
|||||  
Db 3884 ATCGCCCCCGGCAGATGCGGAGCCCGCGGAGGAGATCGCGCGCGCAACAGCACCC 3943  
|||||  
QY 718 CTGAGGAGCAGATCGCTGGATGAGCAGCAACCCCGCGTGGCGTGGCGCCACATCTAC 777  
|||||  
Db 3944 CTGAGGAGCAGATCGCTGGATGAGCAGCAACCCCGCGTGGCGTGGCGCCATCTAC 4003  
|||||  
QY 778 AACCGGTGATCATCTTGGCCTGAACAGATCGTCCGATGTACAGCCCGCTGACATC 837  
|||||  
Db 4004 AACCGGTGATCATCTTGGCCTGAACAGATCGTCCGATGTACAGCCCGCTGACATC 4063  
|||||  
QY 838 CTGACATCGCGCAGGCGCCCAAGAGGCGCTTCCGGAGCTACGTGACCGCTTCTTCAAG 857  
|||||  
Db 4064 CTGACATCGCGCAGGCGCCCAAGAGGCGCTTCCGGAGCTACGTGACCGCTTCTTCAAG 4123  
|||||  
QY 898 ACCCTCGCGCGCAGCAGCCACCAAGGAGCTGAAGAACTGGATGACCGAGACCCGTG 957  
|||||  
Db 4124 ACCCTCGCGCGCAGCAGCAGCCACCAAGGAGCTGAAGAACTGGATGACCGAGACCCGTG 4183  
|||||  
QY 958 GTGCAGAACCCCAACCCCGAC.TGCAACAGCATCTCTCGGGCTCTCGCGCCCGCGCCAC 1017  
|||||  
Db 4184 GTGCAGAACCCCAACCCCGAC.TGCAACAGCATCTCTCGGGCTCTCGCGCCCGCGCCAC 4243  
|||||  
QY 1018 CTGAGGAGATGATACCGCTGCCAGGGCGTGGGGCGCCCGGCCACAGGCGCCCGTG 1077  
|||||  
Db 4244 CTGAGGAGATGATACCGCTGCCAGGGCGTGGGGCGCCCGGCCACAGGCGCCCGTG 4303  
|||||  
QY 1078 CTGGCGGAGCGATGAGCGGAGCCCAAGCGTGAACATCATGATGAGAGAGCAACTTC 1137  
|||||  
Db 4304 CTGGCGGAGCGATGAGCGGAGCCCAAGCGTGAACATCATGATGAGAGAGCAACTTC 4360  
|||||  
QY 1138 AAGGGCCCCGGCGACGCTCAAGTCTCACTGCGGCAAGGAGGCGCCACATCGCCCAAG 1197  
|||||  
Db 4361 AAGGGCCCCGGCGACGCTCAAGTCTCACTGCGGCAAGGAGGCGCCACATCGCCCGC 4420  
|||||  
QY 1198 AACTGCGCGCGCCCGCGCAAGAGAGGCTGCTGAAAGTGGCGAAGGAGGCGCCACAGATG 1257  
|||||  
Db 4421 AACTGCGCGCGCCCGCGCAAGAGAGGCTGCTGAAAGTGGCGAAGGAGGCGCCACAGATG 4480  
|||||  
QY 1258 AAGACTGCAACGAGCGCCAGGCGCAACTCTCTGGGCAAGATCTGGCCGCGCCACAGGGC 1317  
|||||  
Db 4481 AAGACTGCAACGAGCGCCAGGCGCAACTCTCTGGGCAAGATCTGGCCGCGCCACAGGGC 4540  
|||||  
QY 1318 CGCCCGCGCAACTTCTGCGAGAACCGCAGCGAGCGCGCCCGCCCGCCCGTGGCCCGCC 1377  
|||||  
Db 4541 CGCCCGCGCAACTTCTGCGAGCGC-----CCCCGAGCGCCCGCC 4582  
|||||  
QY 1378 CCCCCCGCGGAGAGCTTCGCGCTTCGAGGAGACCAACCCCGCCCGCCCAAGCAGGAGCCCAAG 1437  
|||||  
Db 4583 CCCCCCGCGGAGAGCTTCGCGCTTCGAGGAGACCAACCCCGCCCGCGCAGAGCAGAGCAAG 4642  
|||||  
QY 1438 GACCGGAGCGCTTACCGGAGCGCCCTGACCGCCCTGCGGAGCGCTGTTCGCGAGCGGCC 1497  
|||||  
Db 4643 G-----ACCGGAGAGCCCTGACCGCTGAAGAGCGCTGTTCGCGAGAGCGCC 4650  
|||||  
QY 1498 CTGAGCCAGTAA 1509  
|||||

```
Db      4691 CTGAGCCCAATAA 4702

RESULT 11
US-10-190-435-17
; Sequence 17, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEYER, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTISENSIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PPI8133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 4716
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: GagproteaseRevRef_c
US-10-190-435-17

Query Match      84.5%; Score 1275.2; DB 12; Length 4716;
Best Local Similarity 92.5%; Pred. No. 5.4e-252;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

QY      1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCAGAGCTGGACAAGTGGAGNAGATCGCG 60
Db      1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCAGAGCTGGACAAGTGGAGNAGATCGCG 60
QY      7 ATGGGCGCGCGCGCAGCATCTGCGCGCGCAGAGCTGGACAAGTGGAGNAGATCGCG 66
Db      7 ATGGGCGCGCGCGCAGCATCTGCGCGCGCAGAGCTGGACAAGTGGAGNAGATCGCG 66
QY      61 CTGGCGCGCGCGCGCAGCAAGTACATGCTGAAGCACTCTGGTGTGGCGCGCGCGCGAG 120
Db      61 CTGGCGCGCGCGCGCAGCAAGTACATGCTGAAGCACTCTGGTGTGGCGCGCGCGAG 120
QY      67 CTGGCGCGCGCGCGCAGCAAGTACATGCTGAAGCACTCTGGTGTGGCGCGCGCGAG 126
Db      67 CTGGCGCGCGCGCGCAGCAAGTACATGCTGAAGCACTCTGGTGTGGCGCGCGCGAG 126
QY      121 CTGGAGGCGCTTCGCGCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGCTGCAACAGATC 180
Db      121 CTGGAGGCGCTTCGCGCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGCTGCAACAGATC 180
QY      127 CTGGAGAGTTCGCGCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGCTGCAACAGATC 186
Db      127 CTGGAGAGTTCGCGCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGCTGCAACAGATC 186
QY      181 ATGAAGCACTGCAAGCGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db      181 ATGAAGCACTGCAAGCGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY      187 ATCGCGCAGCTGCAAGCGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 246
Db      187 ATCGCGCAGCTGCAAGCGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 246
QY      241 ACCGTGGCGACCTGTACTGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db      241 ACCGTGGCGACCTGTACTGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY      247 ACCGTGGCGACCTGTACTGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306
Db      247 ACCGTGGCGACCTGTACTGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306
QY      301 CTGGACAAGATGAGAGGAGGAGCAAGTCCGAGGAGAGAGCGCGCGCGCGCGCGCGCGCGCG 360
Db      301 CTGGACAAGATGAGAGGAGGAGCAAGTCCGAGGAGAGAGCGCGCGCGCGCGCGCGCGCGCG 360
QY      307 CTGGACAAGATGAGAGGAGGAGCAAGTCCGAGGAGAGAGCGCGCGCGCGCGCGCGCGCGCG 366
Db      307 CTGGACAAGATGAGAGGAGGAGCAAGTCCGAGGAGAGAGCGCGCGCGCGCGCGCGCGCGCG 366
QY      361 GCGGAC---GGCAAGGTGAGGAGAGAGTACCGCATCGTGCAGAACTTCGAGGGGTGATG 417
Db      361 GCGGAC---GGCAAGGTGAGGAGAGAGTACCGCATCGTGCAGAACTTCGAGGGGTGATG 417
QY      367 GCGGACAAGGGCAAGGTGAGGAGAGAGTACCGCATCGTGCAGAACTTCGAGGGGTGATG 426
Db      367 GCGGACAAGGGCAAGGTGAGGAGAGAGTACCGCATCGTGCAGAACTTCGAGGGGTGATG 426
QY      418 GTGACACCGGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477
Db      418 GTGACACCGGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477
QY      427 GTGACACCGGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486
Db      427 GTGACACCGGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486
QY      478 GCTTCAGCGCGCGAGGTGATCCCATGTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
Db      478 GCTTCAGCGCGCGAGGTGATCCCATGTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
QY      487 GCTTCAGCGCGCGAGGTGATCCCATGTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
Db      487 GCTTCAGCGCGCGAGGTGATCCCATGTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
QY      538 GACCTGAACACGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
Db      538 GACCTGAACACGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
QY      547 GACCTGAACACGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606
Db      547 GACCTGAACACGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606

US-10-190-435-20
; Sequence 20, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEYER, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTISENSIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PPI8133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 4716
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: GagproteaseRevRef_c
US-10-190-435-20

Query Match      84.5%; Score 1275.2; DB 12; Length 4716;
Best Local Similarity 92.5%; Pred. No. 5.4e-252;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

QY      1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCAGAGCTGGACAAGTGGAGNAGATCGCG 60
Db      1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCAGAGCTGGACAAGTGGAGNAGATCGCG 60
QY      7 ATGGGCGCGCGCGCAGCATCTGCGCGCGCAGAGCTGGACAAGTGGAGNAGATCGCG 66
Db      7 ATGGGCGCGCGCGCAGCATCTGCGCGCGCAGAGCTGGACAAGTGGAGNAGATCGCG 66
QY      61 CTGGCGCGCGCGCGCAGCAAGTACATGCTGAAGCACTCTGGTGTGGCGCGCGCGAG 120
Db      61 CTGGCGCGCGCGCGCAGCAAGTACATGCTGAAGCACTCTGGTGTGGCGCGCGCGAG 120
QY      67 CTGGCGCGCGCGCGCAGCAAGTACATGCTGAAGCACTCTGGTGTGGCGCGCGCGAG 126
Db      67 CTGGCGCGCGCGCGCAGCAAGTACATGCTGAAGCACTCTGGTGTGGCGCGCGCGAG 126
QY      121 CTGGAGGCGCTTCGCGCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGCTGCAACAGATC 180
Db      121 CTGGAGGCGCTTCGCGCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGCTGCAACAGATC 180
QY      127 CTGGAGAGTTCGCGCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGCTGCAACAGATC 186
Db      127 CTGGAGAGTTCGCGCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGCTGCAACAGATC 186
QY      181 ATGAAGCACTGCAAGCGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db      181 ATGAAGCACTGCAAGCGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY      187 ATCGCGCAGCTGCAAGCGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 246
Db      187 ATCGCGCAGCTGCAAGCGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 246
QY      241 ACCGTGGCGACCTGTACTGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db      241 ACCGTGGCGACCTGTACTGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY      247 ACCGTGGCGACCTGTACTGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306
Db      247 ACCGTGGCGACCTGTACTGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306
QY      301 CTGGACAAGATGAGAGGAGGAGCAAGTCCGAGGAGAGAGCGCGCGCGCGCGCGCGCGCGCG 360
Db      301 CTGGACAAGATGAGAGGAGGAGCAAGTCCGAGGAGAGAGCGCGCGCGCGCGCGCGCGCGCG 360
QY      307 CTGGACAAGATGAGAGGAGGAGCAAGTCCGAGGAGAGAGCGCGCGCGCGCGCGCGCGCGCG 366
Db      307 CTGGACAAGATGAGAGGAGGAGCAAGTCCGAGGAGAGAGCGCGCGCGCGCGCGCGCGCGCG 366
QY      361 GCGGAC---GGCAAGGTGAGGAGAGAGTACCGCATCGTGCAGAACTTCGAGGGGTGATG 417
Db      361 GCGGAC---GGCAAGGTGAGGAGAGAGTACCGCATCGTGCAGAACTTCGAGGGGTGATG 417
QY      367 GCGGACAAGGGCAAGGTGAGGAGAGAGTACCGCATCGTGCAGAACTTCGAGGGGTGATG 426
Db      367 GCGGACAAGGGCAAGGTGAGGAGAGAGTACCGCATCGTGCAGAACTTCGAGGGGTGATG 426
QY      418 GTGACACCGGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477
Db      418 GTGACACCGGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477
QY      427 GTGACACCGGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486
Db      427 GTGACACCGGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486
QY      478 GCTTCAGCGCGCGAGGTGATCCCATGTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
Db      478 GCTTCAGCGCGCGAGGTGATCCCATGTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
QY      487 GCTTCAGCGCGCGAGGTGATCCCATGTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
Db      487 GCTTCAGCGCGCGAGGTGATCCCATGTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
QY      538 GACCTGAACACGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
Db      538 GACCTGAACACGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
QY      547 GACCTGAACACGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606
Db      547 GACCTGAACACGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606
```

778	AA	CGGTGGATCATCTGGGCTTGACAAAGATCGTGGGATGTACAGCCCGGTGAGCATC	837
787	AA	CGGTGGATCATCTGGGCTTGACAAAGATCGTGGGATGTACAGCCCGGTGAGCATC	846
838	CT	GACATCGCCAGGCGCCCAAGGAGCCCTTCGGGACTACGTGGACCGCTTCTCAAG	897
847	CT	GACATCAAGCAGGGCCCCCAAGAGCCCTTCGGGACTACGTGGACCGCTTCTCAAG	906
898	AC	CTCGGCGCGAGCAGGCCACCCAGGAGCTGAAGAACTGGAATGACCGAGACCTTCTG	957
907	AC	CTCGGCGCGAGCAGGACACCCAGGAGGTGAAGAACTGGAATGACCGAGACCTTCTG	966
958	GT	CAGAAAGCCAAACCCGACTCGAAGACATCTCGGCGCTCTCGGCGCCGCGGCCAAC	1017
967	GT	CAGAAAGCCAAACCCGACTCGAAGACATCTCGGCGCTCTCGGCGCCGCGGCCAAC	1026
1018	CT	GAGGAGATGATACCGCTCTCCAGGCGGTGGGGCGCCCGCCGACAAAGGCCCGGTG	1077
1027	CT	GAGGAGATGATACCGCTCTCCAGGCGGTGGGGCGCCCGCCGACAAAGGCCCGGTG	1086
1078	CT	GCGCGAGCGATGAGCGAGCCACACAGCGTGAACATCATGTGCAAGAGCAACTTC	1137
1087	CT	GCGCGAGCGATGAGCGAGCCACACCGCTGAACATCATGTGCAAGAGCAACTTC	1143
1138	AA	GGGCGCCCGCGCAACGTCAGAGTCTTCAACTGGCGCAAGGAGGCCACATCGCCAA	1197
1144	AA	GGGCGCCCGCGCATCGTCAAGTCTTCAACTGGCGCAAGGAGGCCACATCGCCCGC	1203
1198	AA	CTGGCGCGCCCGCGCAAGAGCGCTGTGAAAGTGGCGCAAGGAGGCCACCATG	1257
1204	AA	CTGGCGCGCCCGCGCAAGAGCGCTGTGAAAGTGGCGCAAGGAGGCCACCATG	1263
1258	AA	GACTGCAACGAGCGCCAGGCGCAACTTCTTGGCAAGATCTTGGCCAGCCCAAGGC	1317
1264	AA	GACTGCAACGAGCGCCAGGCGCAACTTCTTGGCAAGATCTTGGCCAGCCCAAGGC	1323
1318	CG	CGCGCGCAACTTCTTGGCAAGACCGGAGCGCCCGCCCGCCCATCTTGGCCAGCGCC	1377
1324	CG	CGCGCGCAACTTCTTGGCAAGACCGGAGCGCCCGCCCGCCCATCTTGGCCAGCGCC	1385
1378	CC	CGCGCGCGAGAGCTTCCGCTTCGAGGAGACACCCCGCCCGCCCAAGCAGGAGCCCA	1437
1386	CC	CGCGCGCGAGAGCTTCCGCTTCGAGGAGACACCCCGCCCGCCCAAGCAGGAGCCCA	1443
1438	GA	CGCGCGAGCCCTACCGCGAGCCCTGACCGCCCTGCGACACCTTTCGGCAGCGGCC	1497
1426	G	-----ACCGGAGAGCCCTACCGCGAGCCCTGACCGCCCTTTCGGCAGCGGCC	1473
1498	CT	GAGCGACGAA 1509	
1474	CT	GAGCGAGAA 1485	

```

RESULT 13
US-10-190-435-9
; Sequence 9, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBERG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190.435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3930
; TYPE: DNA

```



QY	61	CTCGCCCGCCGGCGCAAGAGCACTACATGCTGAAGCAACCTGGTGTGGGCCAGCCCGCAG	120
DB	67	CTCGCCCGCCGGCGCAAGAGTGCTACATGATGAAGCAACCTGGTGTGGGCCAGCCCGCAG	126
QY	121	CTGGAGGGCTTCGCCCTGTAACCCCGGCGCTGCTGGAGACCGCCGAGGGCTCCAGCAGATC	180
DB	127	CTGGAGAAGTTCCGCTTGAACCCCGGCGCTGCTGGAGACCGAGGGCTTGAAGCAGATC	186
QY	181	ATGAAGCAGCTGCAGCGCCGCTTCAGACCGGACAGCAGAGAGCTGCGCAGCCTGTACAAC	240
DB	187	ATCCGCCAGCTGCACCCCGCTTCAGACCGGACGAGGAGCTGAAGAGCTGTTCAAC	246
QY	241	ACCGTGGCCACCTGTACTGCTGTGACGGCGCGCATGAGGTCCGCGACACCAAGAGGGCC	306
DB	247	ACCGTGGCCACCTGTACTGCTGTGACGAGAAATCGAGTCCGCGACACCAAGAGGGCC	306
QY	301	CTGCACAAGATCGAGGAGGACAGAACAGTCTCCAGCAGCAAGACCCAGTAGGCCAAGGAG	366
DB	307	CTGCACAAGATCGAGGAGGACAGAACAGTGTCCACAGNATCCAGCAGCGCCAGGCC	366
QY	361	GCGCAC---GGCAAGGTGAGCCAGAACTACCCATCGTGCAGAACTGCGAGGCCAGATG	417
DB	367	GCGCACAGGCGAAGGTGAGCCAGAACTACCCATCGTGCAGAACTGCGAGGCCAGATG	426
QY	418	GTGCACAGGCCATCACGCCCGCACGCTCAACGCGTGGTGAAGTGTATGAGGAGAAAG	477
DB	427	GTGCACAGGCCATCACGCCCGCACGCTCAACGCGTGGTGAAGTGTATGAGGAGAAAG	486
QY	478	GCCTTCAGCCCGGAGGTATCCCATGTTTCAGCGGCGCTGAGCGAGGCGCCACCCGCCAG	537
DB	487	GCCTTCAGCCCGGAGGTATCCCATGTTTCAGCGGCGCTGAGCGAGGCGCCACCCGCCAG	546
QY	538	GACCTGAACACGATGTTGAACCGGTGGCGCGCACCCAGACCGCCATGCAAGTGCACAG	597
DB	547	GACCTGAACACGATGTTGAACCGGTGGCGCGCACCCAGACCGCCATGCAAGTGCACAG	606
QY	598	GACACCATCAACAGGAGGAGCGCGAGTGGAGCGGCGTGCACCGCGTGCAGCGCGCGCC	657
DB	607	GACACCATCAACAGGAGGAGCGCGAGTGGAGCGGCGTGCACCGCGTGCAGCGCGCGCC	666
QY	658	GTGCCCGCCGCCAGATGCGCGACCCCGCGGCGACGACATCGCGCGCGCGACCAACCC	717
DB	667	ATGCGCCCGCGCCAGATGCGCGACCCCGCGGAGGACATCGCGCGCGCGACCAACCC	726
QY	718	CTGCAGGAGCAGATCGCTGGATGACACAGCAACCGCCCGTGGCCGTGGCGGACATCTAC	777
DB	727	CTGCAGGAGCAGATCGCTGGATGACACAGCAACCGCCCGTGGCGGAGACATCTAC	786
QY	778	AAGCGGTGGATCTCTCGGCGCTGAACAAGATCGTGGGATGTACAGCCCGGTGAGCATC	837
DB	787	AAGCGGTGGATCTCTCGGCGCTGAACAAGATCGTGGGATGTACAGCCCGGTGAGCATC	846
QY	838	CTGGACATCCGACAGGCCCGCAAGGACCCCTTCGGGACTAGTGGAGCGCTTCTTCAG	897
DB	847	CTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCGGGACTAGTGGAGCGCTTCTTCAG	906
QY	896	ACCTCGCGCGCGAGCAGGCGCACCCAGGACGTGAAGAACTGGATGACCGAGACCCCTGCTG	957
DB	907	ACCTCGCGCGCGAGCAGGCGCACCCAGGAGGTGAAGAACTGGATGACCGAGACCCCTGCTG	966
QY	958	GTGCAGAAGCCCAACCCGAGTGCAGACCATCTCGCGCTCTCGGCCCGCGCGCCACC	1017
DB	967	GTGCAGAAGCCCAACCCGAGTGCAGACCATCTCGCGCTCTCGGCCCGCGCGCCACC	1026
QY	1018	CTGAGGAGATGATGACCGCTGCCAGGGCTGGGCGCGCCCGCCCAACAAGGCCCGGTG	1077
DB	1027	CTGAGGAGATGATGACCGCTGCCAGGGCTGGGCGCGCCCGCCCAACAAGGCCCGGTG	1086
QY	1078	CTGGCCCAAGGGATGAGCCAGGCCCAACAGCGTGAACATCATGATCGAGAGAGCAACTTC	1137
DB	1087	CTGGCCCAAGGGATGAGCCAGGCCCAACAGCGTGAACATCATGATCGAGAGAGCAACTTC	1146

```

RESULT 15
US-10-150-435-11
: Sequence 11, Application US/10190435
: Publication No. US20030143248A1
: GENERAL INFORMATION:
: APPLICANT: ZUR MESEDE, Jan
: APPLICANT: BARNETT, Susan W.
: APPLICANT: LIAN, Ying
: APPLICANT: ENGELBRECHT, Susan
: APPLICANT: VAN RENSBURG, Estrelita J.
: TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
: TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
: FILE REFERENCE: P18133.003 / 2302-18133
: CURRENT APPLICATION NUMBER: US/10/150.435
: CURRENT FILING DATE: 2002-12-30
: NUMBER OF SEQ ID NOS: 319
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 11
: LENGTH: 3930
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: GagcompPolmutina_C
US-10-130-435-11

```

Query March	84.4%	Score 1273.6	DB 12	Length 393C
Best Local Similarity	92.4%	Prod. No. 1.le-251		
Matches 1397; Conservative	0;	Mismatches 79;	Indels 36;	Gaps 4;
Oy	1	ATGGGCGCCGCGCAGCATCTCGGGCGCAGAGCTGGACAGTGGGAGAGATCGC	60	
Db	7	ATGGGCGCCGCGCAGCATCTCGGGCGGCGCAAGCTGGACGCTGGAGCGCATCGC	66	
Oy	61	CTGGCCCCCGCGGAGAAAGCACTACATGCTGAAGCAACCTGGTGTGGGCCAGCGCGAG	120	
Db	67	CTGGCCCCCGCGGCAAGAGTGCTACATGATGAAGCAACCTGGTGTGGGCCAGCGCGAG	126	
Oy	121	CTGGAGGGCTTCGCCCTGAACCCCGGCTGCTGGAGACCGCGGAGGCGCTGCAAGCAGATC	180	
Db	127	CTGGAGAGTTCGCCCTGAACCCCGGCTGCTGAGACCAAGGAGGCTGCAGCAGATC	186	
Oy	181	ATGAAGCAGCTGCAGCGCGCGCTGCAGACCGGCACCGGAGAGCTGCGCAGCGCTGTACAC	240	
Db	187	ATCGCCAGCTGCACCGCGCGCTGCAGACCGGCACCGGAGAGCTGCAAGAGCGCTGTCAAC	246	







Db 361 GACACGGGACACAGCAACAGGTTCAGGCAGAACTACCCCATCGTCAGAACTCCAGGGG 420  
QY 412 CAGATGGTGCACGAGCCATCAGCCCGCCGACCCCTGAACCCCTGGGTGAAGTGA-TCGAS 471  
Db 421 CAGATGGTGCACGAGCCATCTCCCGCCGACCTGAAGCCCTGGGTGAAGTGG-GGAA 480  
QY 472 GAGAGGCGCTTCAGCGCCGAGGTGATCCCATGTTACCGCCCTGAGCGAGGGCGG-CACC 531  
Db 481 GAGAAGGCTTTAGCCCGGAGGTGATACCCATGTTCTCAGCCCTGTCAAGAGGAGCCACC 540  
QY 532 CCCCAGACCTGAACACGATGTTGAACACGCTGGGGGCGCCACACGCGCCCATGCA3ATG 591  
Db 541 CCCCAGATCTGAACACCATGCTCAACACAGTGGGGGACACACGCGCCCATGCA3ATG 600  
QY 592 CTGAAGGACACCATCAAGGAGGAGCGCGGAGTGGGACCGCTGCACCCCTGGTGCAGAGCC 651  
Db 601 CTGAAGGAGACCATCAATGAGGAGGTGCGAATGGGATCGTGTGATCGG-TGCACGA 660  
QY 652 GGGCCGCTGCGCCCGCCGACAGT-GCGCGACCCCGCGGCGAGCAGCATCGCCGCGCCAC 711  
Db 661 GGGCCCATCCACCGGCCAGATGCGTGAGCCAGGGGCTCAGACATCGCCGGAACGACT 720  
QY 712 AGCACCTGACAGGAGCATGCGCTGTGATGACAGCAACCCCGCTGCGCTGGGCGAC 771  
Db 721 AGTACCCCTTCAGGAACAGATCGGCTGGATGACCAAGCCACCCATCCCGTGGGAGAA 780  
QY 772 ATCTAAGCGGTGGATCA-TCCCTGGGCTTGAACAGATCGTGGGATGTACACCGCGTG 831  
Db 781 ATCTAAGCGGTGGATCATCTCTGGGCTTGAACAGATCGTGGGATGTATAGCCCTATC 840  
QY 832 AGCATCTGACATCGCCCGAGGCGCCAGAGGACCTTCGCGACACGTTGAGACACCTTC 891  
Db 841 AGCATCTGACATCGCCCGAGGCGGAGGAACCTTTCGCGACTAGCTGGGACCGTTTC 900  
QY 892 TCAAGACCTGCGCGCGGAGGAGCCACCCAGAGCTGAGAACCTGGATGACACGAGAC 951  
Db 901 TACAAACCGTTCGCGCGGAGGAGGTAGCCAGGAGTGAAGAACCTGGATGACCGAAGC 960  
QY 952 CTGCTGTGACAGAACCCACCCGATGCAAGACCATCTGCGGCTCTGGCGCGCGG 1011  
Db 961 CTGCTGTGACAGAACCCACCGGACTGCAAGACCATCTGAAGCCCTGGCGCGAGG 1020  
QY 1012 GCGACCTGAGGAGATGATGACCGCTGTCAGGGGTGGGCGGCGCGCGCGACAGGCG 1071  
Db 1021 GCTACCTAGAGAAATGATGACCGCTGTCAGGGAGTGGGCGACCCCGCCACAGGGA 1080  
QY 1072 CCGTGTGTCGCGAGCGATGATGACCGCTGTCAGGGGTGGGCGGCGCGCGCGACAG 1128  
Db 1081 CCGGTGTGCTGAGGCGCATGAGCCAGGTGACCACTCCGCTACCATCATGATGACGCG 1140  
QY 1129 AGCAACTCAAGGGCGCCCGCGCGCAACGTCAAATGCTTCAACTCGCGCAAGAGGGCGAC 1188  
Db 1141 GGCAACTTTGGAACCAACACAGATGCTCAAGTCTCACTGTGGCAAGAGAGGGCAC 1200  
QY 1189 ATGCGCANGAACTGCGCGCGCCCGCGCAAGAGGGTGTGTGGAAGTGGCGCAAGAGGGC 1248  
Db 1201 ACAGCCGCAACTGCAAGGGGCGCTAGAAAAAGGGTGTGTGAAATGCGGCAAGGAAGG 1260  
QY 1249 CACAGATGAGGAGCTGCACTGGAGCGCCAGCGCAACTTCCTGGGCRAGATCTGGGCCAGC 1308  
Db 1261 CACAGATGAGGAGCTGTACGGAGACAGGCTTAATTTTATGGGAGATCTGGGCTTCC 1320  
QY 1309 CACAAGGCGCGCCCGCGCAACTTCTCTCAGAAACCGCAGCGAGCGCGCGCCCGCCAGCTG 1368  
Db 1321 TACAAGGAGGCGCAGGAAATTTCTTCAGAGCAG-----ACCGCGC 1362  
QY 1369 CCAACCGCCCGCCCGCGAGAGCTTCGCTGTGAGGAGACACCCCGCCCGCCCAAGCAG 1428  
Db 1363 CCAACCGCCCGCGCGGAGCTTCAGGTCTGGGTGCGGACAAACAACTCCCGCTCCG 1422  
QY 1429 GAGCCCAAGGACCGGAGCCCTACCGCGAGCCCTGACCGCCCTGCGCAGCCTGTTCGC 1488

Db 1423 AAGCAGAGCGCCAGCCCGCAGGGACGGTGTCTTCAACTTCCTCAGGTCAAGCTTTGCG 1482  
QY 1489 AGCGGCCCC 1497  
Db 1483 AGCGACCCC 1491  
RESULTS  
US-09-552-950-5  
Sequence 5, Application US/09552950  
Patent No. 6541248  
GENERAL INFORMATION:  
APPLICANT: Oxford Biomedica (UK) Limited  
TITLE OF INVENTION: Anti-Viral Vectors  
FILE REFERENCE: 674524-2004  
CURRENT APPLICATION NUMBER: US/09/552,950  
CURRENT FILING DATE: 2000-04-20  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 9772  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: pSYNSP  
US-09-552-950-5  
Query Match 61.3%; Score 925.4; DB 4; Length 9772;  
Best Local Similarity 77.3%; Pred. No. 6,8e-128;  
Matches 1175; Conservative 0; Mismatches 316; Indels 30; Gaps 3;  
QY 1 ATGGGCGCGCGCGCGCAGCATCTCGCGCGCGAGAGAGCTGGACAACTGGGAGAGATCCGC 60  
Db 1108 ATGGGCGCGCGCGCGCAGCGTGTGTGCGGCGCGAGCTGGACCGCTGGGAGAGATCCGC 1167  
QY 61 CTGCGCGCGCGCGCGGGAAGAGACATACATGCTGAGAGCACTGCTGTGSGCCAGCCCGAG 120  
Db 1168 CTGCGCGCGCGCGCGGGAAGAGAGTACAAAGCTGAAGCACTGCTGTGSGCCAGCCGAA 1227  
QY 121 CTGCGAGGCTTCGCTGAACCGCGCTGTGGAGAGACCGCGGAGGCTGCAAGCAGATC 180  
Db 1228 CTGCGAGGCTTCGCTGTGAACCCCGGCTTCTGGAGACCGAGCGGCTGCGCGCAGATC 1287  
QY 121 ATGAAGCAGCTGACGCGCGCGCTGCGAGCCGCGACCGAGAGCTGCGCAGCTGTATCAAC 240  
Db 1285 CTGCGCAACTGACGCGCGCGCTGCAACCGCTGCAAGAGAGCTGCGCAGCTGTATCAAC 1347  
QY 241 ACCGTGGCCACCTGTACTGTGCTGACCGCGCGATCGAGTCCGCGACACACAGGAGGCC 300  
Db 1348 ACCGTGGCCACCTGTACTGTGCTGCAACCGCGCATCGAAATCAAGGATACGAAAGAGGCC 1407  
QY 301 CTGGACAAGATCAGGAGGAGCAGCAACAGTCCCGAGCAAGACCCAGCAGCCCAAGGAG 360  
Db 1408 CTGGATAAATCGAGAGGAGCAGCAATAGACCAAAAGAGCCCAACAGCCGCCCGG 1467  
QY 361 GCGCAGG-----CAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCGAGGCC 411  
Db 1468 GACCGGACACAGCAACCCAGCTCAGCCCAAGACTACCCCATCGTCAGAACTCCAGGGG 1527  
QY 412 CAGATGGTGCACGAGCCCATCAGCCCGCCCGCCCATCGAACCCCTGGGTGAGGTGATCGAG 471  
Db 1528 CAGATGGTGCACGAGCCCATCTCCCGCCCGCAGCTGAAGCCCTGGGTGAAGTGTGGAA 1587  
QY 472 GAGAAGCGCTTCAGCGCGCGAGGTGATCCCATGTTCAACCGCTTGAAGCGGCGCCACC 531  
Db 1588 GAGAAGCGCTTTAGCCCGGAGGTGATACCATGTTCTCAGCCCTGTCAAGAGGAGCCACC 1647  
QY 532 CCCCAGGACCTGAACACAGATGTTGAACACCGGTGGGCGGCGCACAGCCCGCCCATGCAGATG 591  
Db 1648 CCCCAGATCTGAACACCATGCTCAACACAGTGGGGGAGACACACAGCCCGCCCATGCAGATG 1707  
QY 592 CTGAAGGACACCATCAACGAGGAGCGCGCGAGTGGACCGCTGCACCCCGCTGACGCCCTGAGGCC 651

[illegible]

### RESULT 3

```

US-09-184-418C-82
; Sequence 82, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1

```



```
Db 1216 GCTGAAGCAATGAGCCAAAGTAACAATACAAACATATCATGCAGAAACCAATTTTAA 1275
QY 1141 GCGCCCGCGGCGACGTCAGTCTTCAACTGCGGCAAGAGGGCCGATGCGCAAGAC 1200
Db 1276 GGCCCTAAAGAATTTGTAATTTTCAACTGCGCAGGGAAGGSCATATAGCCAGAA 1355
QY 1201 TGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCAAGAGGAGGCGCCACCATGAAG 1260
Db 1336 TGCAGAGCTCC1GGGAANAAGCTTTGGAATG:GGAAAGAGAGAGACCAATATGAA 1395
QY 1261 GACTGCACAGAGCGCAGGCAACTTCTGTGGCAAGATCTGGCCAGGCCAAGAGGCGC 1320
Db 1396 GACTGTACTGAGAGACAGGCTTAATTTT1AGGAAATTTGGCCCTCCAGAGAGGAG 1455
QY 1321 CCGCGCAACTTCTGTCAGAACCGCAGCGAGCGCGCCGCGCCACCGTGTCCAGCGCC 1380
Db 1456 CCGGGAAGCTTCTTCAGAACAGACAGAGCGCAACAGCGCCCGCCAGCTCCACAGCC 1515
QY 1381 CCGCGCGAGAGTTCGGCTTCAGGAGAGCACCGCCCGCCCAAGCAGGAGCGCAAGAG 1440
Db 1516 CCAGCAGAGAGCTTCAGTTCAGGAGAGACACCGCTGCGCCGAGCGAGGAGCAGAAG 1575
QY 1441 CCGGAGCCCTACCGGAGCGCCCTGACCGCCCTGCGCAGC 1479
Db 1576 AAGGAACCCCTTAAGTCCCTCAATCACTCTTTGGCAGC 1614

RESULT 5
US-09-184-418C-73
; Sequence 73; Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 73
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate-962M51.8; gene=gag
US-09-184-418C-73

Query Match 47.0%; Score 709.6; Db 4; Length 1485;
Best Local Similarity 59.2%; Pred. No. 2.7e-96;
Matches 100%; Conservative 0; Mismatches 424; Indels 24; Gaps 2;

QY 1 ATGGGCGCCGCGCCGAGCATCTTGGCGCGGAGAACTGAGCAAGTGGAGAGATCCG 65
Db 1 ATGGGTGCGAGAGCGTCAATATTAGAGGGGGAAAAATTAGTAAATCGGAAAAATTAG 65
QY 61 CTSCGCGCCGCGGCAAGAGACTACATGCTGTAAGCACCCTGCTGTGGGCGAGCCGCG 120
Db 61 CTAGGCCAGGGGGAAGAAAGCGCTATATGATAAACACCTTAGTAAGGCAAGCGGAG 120
QY 121 CTGAGAGGCTTCGCCCTGAACCCGCGCTGCTGGAGACCGCGGAGGCTGCAAGAGATC 180
Db 121 CTGGAAGATTTCGCTTAACCTTGCCCTTTAGAACATCACAGGCTGTGAACAAATA 180
QY 181 ATGAAGAGCTGCAGCGCCGCTGACAGCGGACCGAGAGAGCTGCGAGGCTGTACAA 240
Db 181 ATGAACAGCTACAAACAGCTCTTCAGACAGGAACGGAGGAAGCTTAGATCAATATAC 240
QY 241 ACGTGGCCACCTGTACTGCGTGCAGCGCGGCGATCGAGGTCGGGACACCAAGAGCC 300
Db 241 ACAGTAGCAACTCTCTATTGTGTACATGAAGGGGTAGAGGTACGAGACACCAAGAGCC 300
```

```
QY 301 CTGACACAGATCTGAGGAGGAGCAAGTCCCGACGAGAGAGCCACGACGCCA----- 355
Db 301 TTAGACAGATAGAGGAAGAGCAAAACAAATTCAGCAAAAATACGAAAAAACAACAG 360
QY 356 -AGSAGGCGGAGCGCAAGGTGAGCGAGAACTACCCCATCTGTCGCAAACTGCAGGCGCAG 414
Db 361 CAAGCGCTGAGGAAAGGTCAGTCAAAATTAATCTATAGTGCAGAAATCTCCAGGGCAA 420
QY 415 ATGCTGCACAGGCGCATCAGCCCGCGACCCCTGAACGCGCTGGTGAAGGTGATCGAGGAG 474
Db 421 ATGCTACACCAAGAACTATCACTAGAACTTTGAATGCTGGTAAAGTAAAGAGAA 480
QY 475 AAGGCGCTCAGCCCGAGGCTGATCCCATGTCACCGCCCTCAGCGGGCGCCACCCGC 534
Db 481 AAAGCTTTTAGCCCGAGGTAAACCCATGTTTACAGCATTAACAGAGAGGACCCACCA 540
QY 535 CAGGACCTGAACACCATGTTGAACACCGTGGCGGCCACCGCGCCCATGCAGATGCTG 594
Db 541 CAAGATTTAAACACCATGTTAAATACAGTGGGGACATCAAGCAGCCATGCATATGTTA 600
QY 595 AAGCACACATCAACAGGAGGCGCCCGAGTGGGACCCTGTCACCGCCCGTGCAGGCGCGC 654
Db 601 AAAGATACTATCAATGAGGAGGCTGCAGAAATGGGATAGATTACATCCAGTGCATGCG 660
QY 655 CCGTGGCGCCCGGCGCAGATGGCGACCCCGCGGCGAGGACATCGCGCGCCACCCAGC 714
Db 661 CCTATTGCACGAGGCCAATGAGAGACCAAGGGAGTGAATAGCAGGAACACTACTAGT 720
QY 715 ACCCTGCAGGAGCAGATCGCTTGGATGACAGCAACCCCGCGTGCCTGGCGGACATC 774
Db 721 ACCCTCCAAGAACAGATAGCATGATGATGACAAAGTAATCCCGCTATTCAGTGGAGACA 780
QY 775 TACAAGCGGTGGATCATCTGGGCTGACAAAGATGCTGGGATGTACAGCCCGTGAAGC 834
Db 781 TATAAAGATGATTAATCTGGGTTAAATAAATAGTAAGAATGTATAGCCTGCTCAGC 840
QY 835 ATCTGTGACATCCCGCAGGCGCCCAAGGAGCCCTTCGCGCATCTAGTGCAGCGCTTCT 894
Db 841 ATTTGGACATAAACAAAGGCGCAAGGAAACCCCTTAGAGACTATGTAGACGGTCTTC 900
QY 895 AAGACCTTCGCGCGAGCAGCCACCCAGGACGCTGAAGAACTGGATGACGAGACCTGTG 954
Db 901 AAAACTTTAAGAGCTGAACAGGCTACACAAGAAAGTAAAAAATTTGGATGACAGACCT 960
QY 955 CTGCTGCAGACGCCAACCCGACTGCGAGACCATCTCGCGCTCTCGGCCCGCGGCC 1014
Db 961 TTGCTCAAAATGCAACCCAGATTGCAAGAGCATTTTAAAGACATTTAGGACCGAGG 1020
QY 1015 ACCCTGGAGGAGATGATGACCGCTTGGCGCGCTGGCGCGCCCGCGCAAGCGCCGC 1074
Db 1021 ACATTAGAGAAATGATGACAGCTGTCAAGAGTGGGAGGACCTAGCCACAAAGCAAGA 1080
QY 1075 GTGCTGGCGGAGGCGATGAGCAGGCCCAACAGCGTGAACATCATGATGACAGAGAGCA 1134
Db 1081 GTGTGGCTGAGGCAATGAGCAAAACAATAAGTGTAAACATCTGATGCAAGAAAGCAAT 1140
QY 1135 TTCAGGCGCCCGCGCGCAAGCTCAAGTCTTCACTGCGGCAAGGAGGCGCACATCGCC 1194
Db 1141 TTTAAGGAAATAAAGAAATGGTTAAATGTTTTTAAGTGGTAAGGAGGCGCATAGCC 1200
QY 1195 AAGAACTCCCGCGCCCGCGCAAGAGGCTGCTGGAAGTGGCAAGAGGCGCCACCCAG 1254
Db 1201 AGAATTCGAGGCGCCCTAGGAAAGAGGCTGTGGAAATGTGGAAGAGGAGGACACCAA 1260
QY 1255 ATGAAGAGCTGCACGAGCGCGCAGGCCAACTTCCTGGGCAAGATCTGCGCCAGGCAAG 1314
Db 1261 ATGAAGAGCTGTACTGAGAGGCGAGGCTTAATTTT1AGGGAATAATTTGGCTTCC 1320
QY 1315 GCGCGCCCGCGCAACTTCTGTCAGAACCGCAGGAGCGCGCGCCCGCCACCGTGC 1374
Db 1321 GGAAGGCCAGGGAATTTCTCTTCAGAACAG-----GCCAGAGCCCAACA 1362
```

```
Qy 1375 GCGCCCGCCCGCAGAGCTTCGCTTCGAGGAGACACCCCGCCCGCCCAAGAGAGAGGCC 1434
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 1363 GCGCCACCGAGAGAGCTTCAGGTTGCGAGAGACACCCCGCTCGAGAGAGAGAGTGC 1422
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 1435 AAGGACCGCGAGCGCT 1450
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 1423 AAAGACAGGGAAGCCT 1438
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| |

RESULT 6
US-09-184-418C-8
: Sequence 8, Application US/09184418C
: Patent No. 6492110
: GENERAL INFORMATION:
: APPLICANT: Hahn, Beatrice
: APPLICANT: Gao, Feng
: APPLICANT: Shaw, George
: TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
: TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
: FILE REFERENCE: D6287
: CURRENT APPLICATION NUMBER: US/09/184.418C
: CURRENT FILING DATE: 1999-11-02
: NUMBER OF SEQ ID NOS: 112
: SEQ ID NO 8
: LENGTH: 9010
: TYPE: DNA
: ORGANISM: Human immunodeficiency virus type 1
: FEATURE:
: OTHER INFORMATION: Isolate-962M651; 137..1621:"aag"; 2426..4425:"pol";
: OTHER INFORMATION: 4370..4948:"vif"; 4898..5178:"vpr";
: OTHER INFORMATION: 5159..5373-7734:"tat"; 5258..5373-7734..7931:"rev";
: OTHER INFORMATION: 5387..5647:"vif"; 5565..8171:"env"; 8173..8793:"nef";
US-09-184-418C-8

Query Match: 47.08; Score 709.6; DB 4; Length 9010;
Best Local Similarity 69.28; Pred. No. 2.6e-96;
Matches 1008; Conservative 0; Mismatches 424; Indels 24; Gaps 2;

Qy 1 ATGGCGCCCGCCGCGCAGCATCTCTGCGCGGGGAGAGCTGAGCAAGTGGGAGAGATCGCG 60
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 137 ATGGGTGGAGAGAGCTCAAAATTAAGAGGGGGAATAATTAGATAAATGGGAATAATAGG 195
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 61 CTGCGCCCGCGCGGCGAGAGAGCATAGATGCTGAAGCACTGTGTGGGCGAGCGAGAG 120
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 197 CTAAAGGCGAGGGGAGAGAAAGCGCTATATGATAAAGCACTAGTATGGGCAAGCAGGAG 254
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 121 CTGAGGGCTTCGCGCTGAAGCGCGGCTGTGTGAGACCCCGGAGGCTGAGAGCATC 180
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 257 CTGGAAGATTTCGCTTACCTTGGGCTTTTGAAGACATCAGAAAGCTGTGAACAATA 316
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 181 ATGAAGCAGCTGCAGCGCGCCCTGCAGACGGCGCACCGAGGAGCTGCGAGCGCTGTAAAC 240
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 317 ATGAACAGCTACAAACAGCTCTTCAGACAGGAGAGCGAGAACTTAGATCATATACAGC 376
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 241 ACCGTGCGACCCCTGTACTGCTGCAGCGCGGATTCGAGGTTCGCGACACAGAGAGGCC 300
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 377 ACATAGCAACTCTCTATTGTGTACATTAAGGGGTAGAGTACGAGACACCAAGAGAGGCC 436
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 301 CTGCACAGATCGAGAGGAGCAGACAACTCCCGAGCAGAGACCCAGAGGCCA-- 355
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 437 TTACAGAGATAGAGGAGAGCAAAATAATTCAGCAAAAATACAGCAAAAACACAG 496
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 356 -AGGAGCGCGAGCGGAGGAGGAGAGCTACCCCATCGTGCAGAACTTCGAGGGCGCAG 414
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 497 CAAGCGGCTGACGAGGAGGTCAGTCAAAATATCTATAGTGAGAAATCTCCAGGGGCAA 556
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 415 ATGTGTCACAGAGGCCATCAGCCCCCGGCGACCTGAAACGCCCTGGGTGAAGGTATCGAGAG 474
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 557 ATGTGTCACAGAGAACTATCACCTTAGAATTTGAATGCATGCGGTAAAAGTAATAGAGAA 616
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 475 AAGCGCTTCACCCCGGAGGTATCCCATGTTACCGCCCTGAGCGAGGGCGGCCACCC 534
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 617 AAGCGCTTCACCCCGGAGGTATCCCATGTTTACAGCATTTATCAGAGAGGAGGCCACCA 676
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
```

## RESULT 7

```
US-09-184-418C-100
: Sequence 100, Application US/09184418C
: Patent No. 6492110
: GENERAL INFORMATION:
: APPLICANT: Hahn, Beatrice
: APPLICANT: Gao, Feng
```

```
Qy 535 CAGGACCTGAACACGATGTTGAACACGCGTGGCGGCACAGCGCCCATGCGAGATGCTG 594
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 677 CAAGATTAAACACCATGTTAAATCAGTGGGGGACATCAAGCAGCCATCAATATGTA 736
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 595 AAGGACACCATCAAGAGGAGGCGCGAGTGGGACCGCTCCACCCCGTGGAGCGCGCG 654
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 737 AAAGATACTATCAATGAGGAGGCTGAGAAATGGATAGATTACATCCAGTGCATGAGG 796
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 655 CCGTGGCCCGCGGCGAGATGCGGACCCCGCGGCGAGCGAGATGCGCGCGCGACACAG 714
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 797 CTTATTGACCGAGGCCAAATGAGAGAACCAAGGGGAAGTGATAGCAGGAACIACATG 856
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 715 ACCCTGCAGGAGCAGATCGCTTGGATGACCAAGACCCCGCTGCGCTGGCGGACATC 774
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 857 ACCCTCCAGACAGATAGCATGGATGACAAGTAAATCCCTATTCAGTGGGAGACATC 916
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 775 TACAAGCGTGGATCATCTGCGGCTGAAACAGATGTCGCGGATGTACAGCCCGTGAAG 834
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 917 TATAAAGATGGATTAATCTGCGGTTAAATAAATAGTAAGATGTATAGCCCTGTACG 976
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 835 ATCCIGGACATCGCCAGGCGCCCAAGGAGCGCTTCCGCGACTAGTGGACCGCTTCTC 894
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 977 ATTTGGACATAAACAAGGCGCAAGGAACCTTTAGAGACIATGTAGACCGTCTCTC 1036
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 895 AAGACCTCGCGCGGAGGAGCGCCAGGAGCTGGAAGAACTGGATGACCGAGACCGCTG 954
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 1037 AAAACITTTAAGAGCTGAACAGGCTACACAAGTAAGTAAGTAAGTAAGTAAGTAAG 1096
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 955 CTGGTGCAGAGCGCAACCCCGACTGCAGAGCATCTCTCGGCGCTCTCGGCGCGCGCG 1014
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 1097 TTGGTCCAAATGCAAAACAGATTGCAAGAGCATTTTAAAGCATTTAGGACGAGGGGT 1156
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 1015 ACCTGGAGGATGATGACCGCTCCAGGCGCTGGCGCGCGCCGCGCCACAGGCGCGCG 1074
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 1157 ACATTAGAGAAATGATGACGATGTCAGGAGTGGGAGGACCTAGCCCAAGAGCAAG 1216
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 1075 GTCTGGCGCGAGCGGATGAGCCAGGCGCAACAGCTGAACATCATGTGAGAGAGCAAC 1134
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 1217 CTGTTGGCTGAGGCAATGAGCAACAAATAGTGTAAACATCTGATGAGAGAAAGCAAT 1276
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 1135 TTCAAGGGCGCGCGCGAGCTCAAGTCTCACTGCGGCAAGAGGAGGCGGCACATCGCG 1194
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 1277 TTTAAAGGAATAAAGAATGTTAAATGTTTAACTGTGTTAAGGAGGCGACATAGCG 1336
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 1195 AAGAACTGCGCGCGCGCGCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGCGCACAG 1254
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 1337 AGAAATGCGAGGCGCGCTAGCAAAAGGCTGTGGAATGTGGAAGAGGCGACACCA 1396
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 1255 ATGAAGACTGCAGCGGCGCGAGGCGCAACTTCTTCTGGCAAGATCTGGCCCGAGCCACA 1314
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 1397 ATGAAGACTGTACTGAGAGGCGGCTAATTTTATAGGAAATTTGGCTTCCCAAG 1456
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 1315 GCGCGCGCGCGCGCACTCTCTGCAAGACCGAGGAGCGCGCGCGCGCGCGCGCGCGCG 1374
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 1457 GSAAGCGCGCGGAAATTTCTTCCAGAACAG-----GCCAGAGGCCAAG 1498
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 1375 GCGCGCGCGCGCGAGAGCTTCCGCTTCGAGGAGACCAACCCCGCGCGCGCGCGCGCG 1434
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 1499 GCGCGCGCGCGAGAGAGCTTACGTTTCGAGGAGACCAACCCCGCTCGAGGAGGAGTGG 1558
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 1435 AAGGACCGGAGCGCT 1450
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 1559 AAAGACAGGGAAGCCT 1574
    ||||| | | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
```

Query Match	46.9%	Score 708	Pos 4	Length 1475	
Best Local Similarity	59.2%	Prod. No. 4.7e-96			
Matches 1004	Conservative	3	Mismatches 425	Indels 21	Gaps 2
Qy	1	ATGGCGCCCGCCAGCATCTCGCGGGGAGAGTGCACAAAGTGGGAGAGATGCGC	60		
Db	1	ATGGTCCGAGAGCGTCAATATTAAGAGGGGAAATTTAGATACATCGGAAATATCG	60		
Qy	61	CTCGCCCGCGGGCAGAGCACTACATCTGAAGCACTGTGTGTGGCCACCGCAG	120		
Db	61	TTAAGCCAGGGGAAAGAAACATATATATAAACACTTAGTATGGGCAAGCAGAG	120		
Qy	121	CTGAGGGCTTCGCCCTGAACCCCGGCTCTGTGACACCCCGAGGCTTCAAGCAGATC	180		
Db	121	CTGAAAGATTTGGCTTTACCTTGGCTTTTAGAGAGCTCAGACGATGTAAACAATA	180		
Qy	181	ATGAAGAGCTGAGCCCGCCTGCAGACCGGCACCGAGAGCTGCGCAGCTGTACAC	240		
Db	181	ATAAAACNGCTACATCCAGCTCTTAAGACAGAACAGAGAACTTAGTTCATTATCAAC	240		
Qy	241	ACCGTGCCACCCCTGTACTCGCTGCACCCCGGATCGAGGTGCGGACACCAAGAGGCC	300		
Db	241	ACAGTAGCACTCTCTATGTGTACATGCGAGGTAGAGGTACGAGACACCAAGAGGCC	300		
Qy	301	CTGGACAAGATCAGGAGAGCAGACAAGTCTCCAGCAGAGCCACAGAGCCACAGAG	360		
Db	301	TTAGACAAGATAGGAAAGAACAAACAAAGTCAGCAAAATATACAGAGGCAAAAGAG	360		
Qy	361	GCGAGGGCAAGGTGAGCCAGATACCCCATGTGTCAGACCTGCGAGGCGCCAGATGTC	420		
Db	361	GCTGACGGAAGGTCACTCAAAATATTCCTATAGTCAAAATCTCCAAAGGCCAATGCTA	420		
Qy	421	CACAGGCCATACGCCCGGACCCCTGAAGCTGGGTGAAGTGTATCGAGAGAGAGGCC	480		
Db	421	CACAGGCCCTATACCTAGAACTTTGAATGCGTGAAGTAAAGTAAAGTAAAGTAAAG	480		
Qy	481	TTACGCCCGGAGGTGATCCCGCATGTTACCGCCCTGAGCGAGGCGGCACACCCCGAGAC	540		
Db	481	TTTAGCCACAGGTAAACCCATGTTTCACAGCATATACAGAGGAGCCACCGCTCTGAT	540		
Qy	541	CTGNACAGATGTGAACACCGTGGCGGCGCCACGAGCGCCATGCGAGATGCTGAAGGAC	600		
Db	541	TTAAACACCATGTTAAATAGTGGGGGACATCAGCAGCCATGCAAAATGTTAAAGAT	600		
Qy	601	ACCATCAACAGGAGCGCGGAGTGGAGCGCTGCAACCCCGTGCAGGCGGGCGCGGTG	660		
Db	601	ACCATCAACAGGAGCGTGGGATGGATAGATTACATCCAGTACATGCGAGGCCATAT	660		
Qy	661	GCCCCCGGCAGATGCGGAGACCCCGCGGAGCAGACATCGCGCGGCCACCGAGACCCCTG	720		
Db	661	CCACAGGCCCATGAGAGAACCAAGGGGAGTGTATAGCAGGAAGTACTAGTACCGCTT	720		
Qy	721	CAGGACAGATCGCTGATGATGACCAACCCCGCTGCGCGTGGCGGACATCTACAAG	780		
Db	721	CAGGACAAATAGCATGGATGACAGGTAAACCCACTATTTCAGTGGGAGACATCTATAA	780		
Qy	781	CGGTGGATCATCTCTGGCCCTGAACAAGATCTGTGCGGATGTACAGCCCGGTGAGCATCTG	840		

```

1156 GAAGAAAGTGTACAGCATGTCTCAAGGAGTGGGAGGACCTAGCCACAAAGCAAGAGTGTG 1217
1061 GCGGAGCGGATAGCGAGGCGCAACAGCGTCAACATATGATGCAGAAAGAGCAACTTCAAG 1140
1218 GCTGAGCGCATGAGCCCATCATAG---TACATATATGTCCAGAGAGCAATTTTAAA 1274
1141 GGGCCCCGGCGCAACCTCAAGTCTTCACTGTGGCAAGGAGGCGCCACATGCGCAAGAAC 1200
1275 GGCCCTAAAGAAATTGTTAAATGCTTCACTGTGGCAAGGAGGCGCATAGCCAGAAAT 1334
1201 TGGCGGCGCCCCCGCAGAAAGGGCTGCTGAAGTTCGGCAGGAGGCGCCACCAGATGAAG 1260
1335 TGCAGGCGCCCTAGAAAGAGGCTGTGTGAAATGTGGCAAGAGACACCAAAATGAAA 1394
1261 GACTGCACCCAGCGCCAGGCGCAACTTCTCTGGCAAGATCTGGCCAGCCACAAAGGCGCG 1320
1395 GACGTACTGAGAGCGAGGCTAATTTTATAGGAAATTTGGCGCTTCCACAGGGGAGG 1454
1321 CCGCGCAACTTCCTGCAGAACCGCAGCGCCCGCCCGCCCGCCACCGTGCACCACCGCCCC 1380
1455 CCAGGGAATTCCTTCAAAACAG-----GCAGAGCCAAACAGCCCGCA 1496
1381 CCGCGCGAGAGCTTCGCTTCGAGGACACACCCCGCCCGCCCGCCCGCCAGCGCCAGGAC 1440
1497 CCAGCAGAGAGCTTCAGGTTCAAGGAGACAAACCCCGCTCCGAGAGAGAGTCGAAAGAC 1556
1441 CGCGAGCGCCT 1450
1557 AGGAAACCT 1566

RESULT 9
CS-09-184-418C-38
; Sequence 38, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184.418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 38
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=92RW009; gene=gag
US-09-184-418C-38

Query Match 44.0%; Score 663.6; DR 4; Length 1486;
Best Local Similarity 68.8%; Pred. No. 1.5e-89;
Matches 942; Conservative 0; Mismatches 424; Indels 4; Gaps 2;

QY 1 ATGGGCGCCGCGCCACCATCTTCGCGGCGGAGAGCTGGACAAGTGGGAGAAGA-TCCG 59
DB 1 ATGGGTGCGAGCGCTCAATATTAAAGGCGGAAAATTAGATGCTGGGAAAAAATTAAG 60
QY 60 CCGTGGCGCCCGCGGCAAGCACTACATGCTGAAGCAGCTGTGTGGGCGACCGCGCA 119
DB 61 TTAAGCCAGGGGGGAAGAAACATATATGATGAACACCTAGTAGTGGCAGCAGGGA 120
QY 120 GCTGGAGGGCTTCGCCCTGAACCCCGGCCCTGCTGGAGACCGCGCGAGGCGTGCACAGCAGAT 179
DB 121 GCTGGAAGAGATTGCACTTAAACCTGACCTTTTAGAGACACCAAGAGGCTGTAAACAAAT 180
QY 180 CATGAAGCAGCTGAGCCCGCCCTGAGACCGGCGCACCGAGGAGCTGCCAGCCTGTACAA 239
DB 181 AATGAGACAGCTGCAACCGAGCTTTCAGAGAGGAAACAGATGAACCTTAGGTCATTATATA 240

```



```
QY 240 CACGTCGCCACCCCTGTACTGCGTGCACGCCGCGCAUCCAGGTCGGGACACCAAGAGGC 259
DB 241 TAGATGACCAACCTCTATTGTGTACATCAAAAGATAGATGTAAAGACACCAAGAGGC 350
QY 300 CCGTGCAAGATCGAGGAGGAGCAACAAGTCCCGACAGAACACCCAGCAGCCCAAGA 359
DB 301 CTTAGACAAGATAGAGGAGCAACAACAAGTCCAGCAAAAACACAGCAGCAGCAAGC 360
QY 360 GCGCGAC---GGCAAGTGTAGCAGAACTACCCCATGTCAGAACTGTGAGGGGAGAT 416
DB 361 AGCTGACAAGAGAAAGTCACTCAAAATACCCCTATAGTCAAAATTCACAAAGGCAAT 420
QY 417 GGTGCACAGCCATCAGCCCGCCACCCCTGAAGCGTGGTGAAGGTGATCAGAGAA 476
DB 421 GGTACACAGGCCATATCACCTAGAACTTTGAATGCGTGGGTAAAGTAAATAGAGAA 480
QY 477 GGCCTTCAGCCCGGAGGTATGCCCATGTTCCAGCGCCTGAGGAGGGCCGCCACCCCA 536
DB 481 GGCCTTTAGCCAGAGGTAATACCCATGTTTACAGCAATTATCAGAAAGGACCCGCCACA 540
QY 537 GGACCTCAACACGATGTTGACACCGTGGGCGCCACAGCGCCCATCGAGATGCTGAA 596
DB 541 AGATTAAACACCAATGCTAAATACAGTGGGGGACATCAAGCAGCCATCAAAATGTAAA 600
QY 597 GGACACCATCAACGAGGAGCGCCCGCAGTGGGAGCGGCTGCACCCCTGCGAGGCGCGGC 656
DB 601 AGATACATCAATGAGGAGCGTCGAGAGTGGGTATAGGTATCATCCATGCGAGCGGSC 660
QY 657 CGTGGCCCCCGGCAGATGUCGCAACCCCGCGGCGAGCATGCGCGGCCACACAGCAC 716
DB 661 TGTTCGCCAGCCAGATAAGAGAAACCAAGGGGAAGTGACATAGCAGGAACACTAGTAG 720
QY 717 CCGTGACAGGACAGATCGCTGGATGACCGACGACCCCGCGTCCCGTGGGACATCTA 776
DB 721 CCTTCAGGACAAATAGCATGGATGACAAATACCCACCTATTCCAGTGGGAGAAATTA 780
QY 777 CAAGCGGTGGATCATCTCGGCGCTGAACAAGATCGTGGGATGTACAGCCCGGTGACAT 836
DB 781 TAAAGATGGATATCTCGGGTTAATAAATAGTAAGATGTATAGCCCTGTCAACAT 840
QY 837 CCGGACATCGCCAGGCGCCCAAGGAGCCCTTCGCGGCTACTGTGACCGCTTCCTCAA 896
DB 841 ATTGGACATAAAGCAAGGCGCCCAAGGAGCCCTTTAGAGACTATGTACACCGGTTCTTAA 900
QY 897 GACCTCGCGCCAGCAGGCGCACCCAGGAGTGAAGACTTGGATGACCGAGACCCCTGCT 956
DB 901 AACCTTAAGCCCGACAAAGCTTCACAGATGTAAATAATTTGAATGACATAGCTTGT 960
QY 957 GGTGCAAGCGCAACCCCGACTGCAAGACCATCTCGCGGCTCTCGGCGCCCGCGCCAC 1016
DB 961 AGTCCAAATGCGAAGCCAGATGTAAGACCAATTTAAGACCAATTTAGGCGCTAGGGGCTC 1020
QY 1017 CCGGAGGAGATGATACCCCTGCCAGCGGCTGGGCGGCCCGCCCGCCACAGCCCGCT 1076
DB 1021 ATTGAAGAAATATGACAGCATCCAGGAGTGGAGGACCCCGCCATTAAGCAAGCGGT 1080
QY 1077 CCGCGCAGCGGATGAGCGAGGCGCAACAGGCTGAACATCATGATGAGAGAGCAACTT 1136
DB 1081 TTTGGCTGAACCATATGACCCAAATGACAAACCAACAAATATATGATGACAGAGCAAT 1140
QY 1137 CAAGGGCCCCCGGCAACGCTCAAGTGTCTTCACTGGGCGAAGGAGGCGCCACATCGCAA 1196
DB 1141 TAAGGGCCAGAGAGAAATATTAGTGTTCACCTGTGGCAAGAGAGACACCTAGCCAG 1200
QY 1197 GAATGCGCGCCCGCCCGCAAGAGAGGCTGTGGAATGCGGCAAGGAGGCGCCACAGAT 1256
DB 1201 AAATTGCGAGGCGCCCTAGAAAAAGGGCTGTGGAATGTGGAAGGAGGAGGACACCAAT 1260
QY 1257 GAAGGACTGACCGAGCGCCAGGCGCACTTCTGGGCAAGATCTGGGCCAGCCACAGGG 1316
DB 1261 GAAAGACTGCACTGAGAGACAGGCTAAATTTTATAGGGAATTTTGGGCTTCCACAAAGG 1320
QY 1317 CCGCCCCCGCACTTCTGTCAGAACCCGCGAGGCGCCCGCCCGCCCGCCCGCC 1366
```

DB 1321 GAGCCAGGAAATTTCCCCAGAGCAGACTGGAGCCACAGCCCCACCAS 1370

## RESULT 10

US-09-184-418C-4

: Sequence 4, Application US/09184418C

: Patent No. 6492110

: GENERAL INFORMATION:

: APPLICANT: Hahn, Beatrice

: APPLICANT: Gao, Feng

: APPLICANT: Shaw, George

: TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN

: TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1

: FILE REFERENCE: D6287

: CURRENT APPLICATION NUMBER: US/09/184,418C

: CURRENT FILING DATE: 1999-11-02

: NUMBER OF SEQ ID NOS: 112

: SEQ ID NO 4

: LENGTH: 8992

: TYPE: DNA

: ORGANISM: Human immunodeficiency virus type 1

: FEATURE:

: OTHER INFORMATION: isolate=92RW009; 139.1624:gaq; 1690.4428:pol(N-terminus unce

: OTHER INFORMATION: 4373.4951:vif; 4891.5181:vpr; 5162.7801:tat; 5301.7958:rev;

: OTHER INFORMATION: 5403.5648:vpu; 5566.8148:env; 8150.8773:nef

US-09-184-418C-4

Query Match: 44.0%; Score 663.6; DB 4; Length 8992;

Best Local Similarity 68.8%; Pred. No. 1.4e-89;

Matches 942; Conservative 0; Mismatches 424; Indels 4; Gaps 2;

QY 1 ATGGCGCGCCCGCCAGCATCTCGCGGCGGAGAACTGGACAAGTGGGAGAGA-TCCG 59

DB 139 ATGGGTGCGAGAGCGTCAATATTAAGAGCGGAAAAATTATGCTCGTGGGAAAAAATTAG 198

QY 60 CCGTGGCGCCCGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCACGCCGA 119

DB 199 TTAAGCCAGGGGGGAAAGAAACATATATGATGAACACCTAGTATGGCAAGCAGGA 258

QY 120 GCTGGAGGGTTCGCGCTGAACCCCGCCCTGTGGAGACCGCGGAGGTGCAAGCAGAT 179

DB 259 GCTGGAAGATTTGCACTTTAACCTTGACCTTTTAGAGACACCAAGAGGCTGTAACAAAT 318

QY 180 CATGAGACGCTCAGCGCCCGCGGAGCGCCGAGAGGAGCTGCGAGCGCTGTACAA 239

DB 319 AATGACAGCTGCACACCGAGCTTCAGACAGAAACAGATGAACCTTAGGTCAATATAA 378

QY 240 CACCGTGGCCACCTGTACTGCTGCACCGCGGCATCGAGGTCCGCGACACCAAGAGGC 299

DB 379 TACAGTAGCAACCTCTATTGTGTACATCAAAAGATAGATGTAAGACACCAAGAGGC 438

QY 300 CCGTGGCAAGATCGAGGAGGAGAGAACAGTCCAGCAGAACCCAGCAGGCCAGGA 359

DB 439 CTTAGACAGATAGAGAGAAACAAAGTTCAGCAAAAACACACAGCAGCGCAAGC 498

QY 360 GCGCGAC---GGCAAGTGTAGCGCAACTACCCCATCTGTCGAGAACTGTCAGCGCCAGAT 416

DB 499 AGCTGACAAAGGAAAGTCACTCAAAATTTACCTTAGTGTCAAAATGCACAGGCAAT 558

QY 417 GGTGCACAGGCCATCAGCCCGCCAGCCCTGAACGCTGGGTGAAGGTGATCAGAGAA 476

DB 559 GGTACACAGGCCATATCACCTAGAACTTTGAATGCGTGGGTAAAGTAAATAGAGAA 618

QY 477 GGCCTTCAGCCCGGAGGTATCCCATGTTTCACCGCGCTGAGCGAGGCGCCACCCCA 536

DB 619 GGCCTTTAGCAAGAGGTAAATACCCATGTTTACAGCATTTATCAGAAAGGAGCCACCA 678

QY 537 GGACCTGAACACGATGTTGAACACCGTGGCGCGCCAGCCAGCCCATGCAGATGCTGAA 596

DB 679 AGATTTAAACACCATGCTAAATACAGTGGGGGACATCAAGACAGCCATGCCAATGTAAA 738

QY 597 GGACACCATCAAGCAGGAGGCGCCCGAGTGGGACCGCTGCACCCCGTGCAGCGCCGCC 656





```

1198 TGAGGGCCCCAGAAAAGGCGTGTGGAAGTGTGGAAGAGGGGACACCAATGAG 1257
1261 GACTGACCGAGCGGCGAGCCAACTTCCTGGGCAAGATCTGGCCAGACCAAGGGGCGG 1320
1258 GACTGCACTGAGAGAGGCTTAATTTTGGGAAATTTGGCTTCCAAAGGGGAGG 1317
1321 CCGGGCACTTCTGCGAAGCGGAGCGAGCCGCGCGCCCGCCACCGTGCACCGCCGCC 1350
1318 CCGGGAACCTTCATCCAGAACAG- - - - -GCCAGAGCCGTCAGCCCGG 1359
1381 CCGCGCGAGAGCTTCGCTTCAGAGAGACACCCCGCCCGCCCGAGCAGGAGCCCAAGGAC 1440
1360 CCAGCAGAGAGCTTCAGTTTCGGGAGGAGAGACACCCCATTCGAGACGAGCAGNAA 1419
1441 CCGGAGCCCTACCGGAGCGCCCTGACCGCCCTGCGCAGCCTGTTGGGAGGGGCC 1497
1420 GAGGAGGAGCTGACCTCCCTTAGCTTCCCTCAAATCACTCTTTGGCAAGGAGCC 1476

```

## RESULT 13

US-08-418-848A-9

; Sequence 9, Application US/08418848A

; Patent No. 5847096

; GENERAL INFORMATION:

; APPLICANT: SCHUBERT, MANFRED, HARMISON II,

; APPLICANT: GEORGE G., CHANG-JIE, CHEN, BANJERJEA, AKHIL

; TITLE OF INVENTION: DEFECTIVE, INTERFERING

; TITLE OF INVENTION: HIV PARTICLES

; NUMBER OF SEQUENCES: 77

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN &amp; F'NNEGAN, L.L.P.

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; COUNTRY: U.S.A.

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORD PERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/4-8,848A

; FILING DATE: 07-APR-1995

; CLASSIFICATION: 526

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/936,849

; FILING DATE: 28-AUG-1992

; CLASSIFICATION: 526

; ATTORNEY/AGENT INFORMATION:

; NAME: RICHARD W. BORK

; REGISTRATION NUMBER: 36,459

; REFERENCE/DOCKET NUMBER: 2026-4091US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-758-4800

; TELEFAX: 212-751-6849

; TELEX: 42-792

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7399 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-418-848A-9

Query Match

Best Local Similarity 42.6%; Score 642.2; DB 2; Length 7399;

Matches 998; Conservative 0; Mismatches 493; Indels 30; Gaps 3;

```

1 ATGGGCGCCGCGCCAGCATCTCGCGCGGAGAGACCTGGACAGCTGGGAGAAATCCG 60
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

751 ATGGGTGCGAGAGCGTCGTATTAAAGCGGGGAGAGATTAGATAAATGGGAAAAATTCGG 810
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

61 CTGCGCCCGCGGCGGCAAGAGCACTACATGCTGGAAGCACCTGCTGTGGGCCAGCGCGAG 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
811 TTAAGCCAGGGGAAAGAAACAATATAAATATAAACAATATAGTATGGCAAGCAGGAG 870
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 CTGGAGGGCTTCGGCTTGACCCCGCGCTGTGGAGAGCCGCGGAGGCTTGAAGCAGATC 180
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
871 CTAGAAGGATTGCGAGTTAATCTTGGCTTTTAGAGACATCAGAAGGCTGTAGACAAATA 930
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 ATGAAGACCTGCAGTCCGCTTCAGAGCGGACCGGAGAGCTCCGAGGCTGTATACAA 240
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
931 CTGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATATATAAT 990
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 ACCGTGGCCACCTGTACTGCTGCACCGCGGATCAGAGTCCGCGACCAAGAGGCC 300
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
991 ACAATAGCAGTCTCTTATTGTGTGCATCAAGGATAGATGTAAAAAGACCAAGAGGCC 1050
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 CTGCAACAAGATCGAGGAGGAGCAAGTCCCGCAGAGAGACCCAGCAGCG- - - - - 353
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1051 TTAGATAAGATAGAGGAGGAGCAAAAGTAAAGAAAAGGACACAGCAGCAGCT 1110
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
354 --CAAGGAGCGCGGCAAGGTGAGCCAGAACTACCCATCGTGCAGAACTCGAGGCC 411
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1111 GACACAGAAACAACAGCCAGGTGAGCCAAAATTACCCTATAGTCGCAAGCTCCAGGG 1170
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
412 CAGATGTCACAGGCCCATCAGCCCGCCAGCCCTGAAGCTGGGTGAAGGTATCAG 471
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1171 CAAATGTCATCAGGCCATATCACCTAGAACITTTAAATGATGGTAAAGTAGTAGAA 1230
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
472 GAGAAGGCTTCAGCCCGAGGTGATCCCATGTTACCGGCTTACAGGCGCTGAGGAGGCCACC 531
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1231 GAGAAGGCTTCAGCCCGAGGATATACCATGTTTTCAGCATTTATCAGAAAGAGGCCACC 1290
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
532 CCCAGGAGCTGAACACAGTGTGAAACCGTGGCGGCGCACAGCGCCGATCGAGATG 591
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1291 CCACAGATTTAAATACCATGCTAAACAGTGGGGGACATCAAGACGACCATGCAATG 1350
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
592 CTGAAGGACACCATCAACGAGGAGCGCGGAGTGGAGCGCTGCACCGCTGGAGGCC 651
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1351 TTAAGAGAGACCATCAATGAGGAGTGCAGATGGGATAGATTCATCCATCCAGTGCATGA 1410
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
652 GGCCCGTGGCCCGGCGAGATGGCGACCCCGCGGCGGAGATCGCGGCGGCCACC 711
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1411 GGCCCTATTGCACGAGCGGATGAGAGAACCAAGGAGAGTACATAGCAGGAACACTACT 1470
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
712 AGCACCTGCAAGAGAGATCGCTGGATGATGATGATGATGATGATGATGATGATGATGATG 771
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1471 AGTACCTTCAGGAAACAATAGGATGATGATGATGATGATGATGATGATGATGATGATG 1530
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
772 ATCTAAGGCGGTGATCATCTGCGCTGACACAGATGCTGCGGATGTACAGCCCGCTG 831
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1531 ATCTATAAGAGATGGATAATCTGGGATTAATAAATAGTAGTAAGTATGTATAGCCCTACC 1590
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
932 AGCATCTGGACATCGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 891
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1591 AGCATCTGGACATGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1650
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
892 TTCAGACCGCTGCGCGCGGAGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 951
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1651 TATAAAACTCTTAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1710
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
952 CTGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1011
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1711 TTCTTGGTCCAAATCGGAACCCAGATTTGTAAGATATTTTAAAGCAATTTGGGACGAGGA 1770
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1012 GCCACCTGAGAGGAGATGATGAGCGGCTGCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1071
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1771 GCGACACTAGAGAGATGATGACAGCATGTACAGGAGTGGGGGAGCGGCGGCGGCGGCGGCGG 1830
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1072 CCGGTGCTGGCGAGGCGGATGAGCGGAGG- - -CCACAGCGGTGAACATCATCATGATGAGAG 1128
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1831 AGAGTTTGGCTGAAGCAATGAGCCAAAGTAAACAAATCCAGCTACCATAATGATACAGAAA 1890
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1129 AGCAACTTCAAGGGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1188
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

1891 3GCATTTTGGGACCAAGAGAGACTGTAAAGTGTTCATTTGGCCAGAGAGCCAC 1950
1189 ATGCCCAAGACTGCGGCCGCCGCCCAAGAGAGTGTGCTGAAGTGGGGCAAGAGAGGC 1248
1951 ATAGCCAAAATTTGAGAGGCCCTAGGAAAAGGGCTGTGGAAATGTGGAAGGAAGA 2010
1249 CACCAGATGAAGACTGCAAGGAGGCCAGGCACTTCCCTGGCAAGATCTGGCCAGC 1308
2011 CACCAATGAAGATGTACTGAGACAGAGGCTAAATTTTAGGAAGATC:GAGCTTCC 2070
1309 CACAAGGGCGGCCGCCGCCCAAGTCTCTGCAAGAGAGAGAGAGAGAGAGAGAGAG 1368
2071 CACAAGGAAGGCCAGGGAATTTCTTCAGAGCAG-----ACCAGAG 2112
1369 CCGACCGCGCGCGCGCGCGAGAGCTTCCGCTTGGAGGACACACCGCGCGCGCGAGCAG 1428
2113 CCACAGCGCCAGCAGAGAGAGCTTCAGGTTTGGGGAAGAGACACAGCTCCTCAG 2172
1429 GAGCCCAAGGAGCGGAGCGCTACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1488
2173 AAGCAGAGCGGATAGCAGAGAGCTGTATCGTTTACCTTCCCTCAGATCACTCTTGC 2232
1489 AGCGGCGCGCGCTGAGCGAG:AA 1509
2233 AGCGACCGCTCGTCACAATAA 2253

```

## RESULT 14

```

US-08-188-583-5
: Sequence 5, Application US/08188583
: Patent No. 5851813
: GENERAL INFORMATION:
: APPLICANT: Desrosiers, Ronald C.
: TITLE OF INVENTION: PRIMATE LENTIVIRUS VACCINES
: NUMBER OF SEQUENCES: 57
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: OPERATING SYSTEM: IBM PS/2 Model 507 or 555X
: SOFTWARE: WordPerfect (Version 5.0)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/188,583
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/727,494
: FILING DATE: July 9, 1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/551,945
: FILING DATE: July 12, 1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Freeman, John W.
: REGISTRATION NUMBER: Req. No. 5851813 29,066
: REFERENCE/DOCKET NUMBER: 00246/079002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 542-5070
: TELEFAX: (617) 542-8906
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 9709
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-188-583-5

```

```

Query Match 42.6%; Score 642.2; DB 2; Length 9709;
Best Local Similarity 65.6%; Pred. No. 1.9e-86;
Matches 958; Conservative 0; Mismatches 493; Indels 30; Gaps 4;
QY 1 ATGGCGCGCGCGCGAGCATCTCGCGCGCGAGAGCTGGACAAAGTGGGAGAGATCCGC 60
DB 750 ATGGGTGGAGCGTGGGTATTAAGCGGGGGAGATTAAGATAATGGGAAAAATTCGG 849
QY 51 CTGCGCGCGCGCGCGAGAGACATACATGCTGAAGCACCTGGTGTGGGCGAGCGCGAG 120
DB 850 TTAAGCGCGAGGGGAAAGAAACAA:ATAAACTAAACATATAGTAGTGGCAACGAGGAG 909
QY 121 CTGAGGCGCTTCGCCCTCAACCGCGCTGCTGAGACCGCGGAGGCTGCAACGACATC 180
DB 910 CTAGACGAATTCGAGGTTAATCTCGGCTTTAGAGACATCAGAGAGCTGTAGACAATA 969
QY 181 ATGAGGAGCTGCGAGCGCGCGCTGAGACCGCGAGGAGCTGCGCAGCGCTGACAAAC 240
DB 970 CTGCGACAGCTACACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATATATAAT 1029
QY 241 ACCGTGGCGACCCCTGTACTGCTGACGCGCGGATCGAGGTCCGCGACACCAAGAGGCC 308
DB 1030 ACAATAGCAGTCTCTATTGTGTGTCATCAAGAGATAGATGTAAAAGACACCAAGSAGGCC 1089
QY 301 CTGGACAAAGATCGAGGAGGAGCAGACAAAGTCCCGAGGAGAGACCCAGCGG----- 353
DB 1090 TTAGATAAGTAGAGGAAGAGCNAACAAAAGTAAAGAAAGGCAAGCAGAGCAGCT 1149
QY 354 --CAAGGAGCGCGAGCGCAAGGTGAGCCAGAACTACCCATCGTCCAGAACTCGAGGC 411
DB 1150 GACACAGGAAACACAGCGCGAGTCAAGCAAAATACCTATAGTGCAGAACTCCAGGGG 1209
QY 412 CAGATGTTGACACAGGCGCATCAGCGCGCGCGACCTGAAGCGCTGGTGAAGGTGATCAG 471
DB 1210 CAAATGGTACATCAGGCGCATACCTAGAACITTTAAATGCGATGGTAAAAAGTASTAGAA 1269
QY 472 GAGAAGGCTTCAGCGCGCGGCTGATCCCATGTTCAAGCGCGCTGAGCGGGCGGCACC 531
DB 1270 GAGAAGGCTTCAGCGCGCGAGTAATACCATGTTTCAGCATTTATCAGAAGGAGCCACC 1329
QY 532 CCGCAGGACCTGAACACGATGTTGAACACCGTGGCGGCGCCACCGAGCGCGCATCAGATG 591
DB 1330 CCACAAGATTTAAATACCATACCTAAACACAGCTGGGGGAGACATCAAGCAGCATGCAATG 1389
QY 592 CTGAGGACACCATCAACAGGAGGCGCGCGAGTGGGACCGCTGACCGCTGCAGCGCC 651
DB 1390 TTAAGAGAGACCATCAATGAGGAGG:GCAGAA:GGGATAGTTGATCCATCCAGTCATGCA 1449
QY 652 GCGCCCGTGGCGCGCGCGAGATGCGCGACCGCGCGCGCGCGCGCATGCGCGCGCGCAC 712
DB 1450 GGGCGCTATTGACCGCGCGCATGAGAGNACCAAGGCGAAGTGACATAGCAGAACTACT 1509
QY 712 AGCACCTTCAGGAGCGAGATCGCCCGGATGACAGCAACCGCGCGCGCGCGCGCGAGAC 771
DB 1510 AGTACCGCTTCAGGACAAATAGGATGGATGATGACACATAATCCACCTATCCCAAGGAA 1569
QY 772 ATCTACAAGCGGTGATCATCTCGGCGCTGAACAGATGCTGGCGATGTACAGCGCGGTG 831
DB 1570 ATCTATAAAGATGATTAATCTCGGATTAATAAATAGTAAGAATGTATAGCGCTTACC 1629
QY 832 AGCATCTCGGACATCGCGCGAGGCGCGCGAGAGCGCTTCCGCGAGTACGTGGAAGCGCTTC 891
DB 1630 AGCATCTCGGACATGAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1689
QY 892 TTCAGACCGCTTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 951
DB 1690 TATAAAGCTTAAGAGCGCGAGCAAGCTTCACAAAGAGGTAATAAATGATGAGAGAAACC 1749
QY 952 CTGCTGTGCGAGAGCGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1011
DB 1750 TTGTTGGTCCAAATGCGGACCCAGATTTGAAGACTATTTTAAAGCATTTGGGACCGAG 1809

```

```
QY 1012 GCCACCTTGGAGGATGATGACCGCCCTGCGAGGCGCTGGCGCGCCCGGCCACAGAGGCC 1071
Db 1810 GCGACAGTAGAAGAAATGATCAGACATGTCAGGGAGTGGGGGACCGCGCCATAAAGCA 1869
QY 1072 CGCTGTGTCGGCGGAGGATGAGCAGG---CCACAGCGGTGAACATCATGTCAGAG 1128
Db 1870 AGAGTTTGGGTGAAGCAATGAGCAAGTAACAAATCCAGCTACCATTAATGATACAGAA 1929
QY 1129 AGCAACTTCAAGGCCCCCGCGGCAACGTCGAAGTGCCTTCAACTGCGGCGCAAGGGCCAG 1188
Db 1930 GGCRAITTTAGAACCAAGAAAGACTGTTAAGTGTTCAATTGTCGCAAGAGGGCCAG 1989
QY 1189 ATCGCCAAAGAACTGCGCGCCCCCGGCAAGAGGCTGCTGGAAGTGGCGCAAGCAAGGC 1248
Db 1990 ATAGCCAAAAATTCAGGCGCCCTAGGAAAAGGGCTGTTGGAATGTGGAAGGAAGGA 2049
QY 1249 CACAGATGAAGGATGACCGGAGCGGCGCAGCGCCACTTCCTGGGCAAGATCTGCCCGAG 1309
Db 2050 CACCAATGAAGATGTACTGAGAGCAGGCTAATTTTTAGGGAAGATCTGGCCTTCC 2109
QY 1309 CACAAGGCGCGCGCGGCAACTTCTGCGAAGCCGCGCAGCGCGCGCGCGCGCGCGCGTG 1368
Db 2110 CACAAGGGAAGGCCAGGGAATTTCTTCAGAGCAG-----ACCAGAG 2151
QY 1369 CCCACCGCGCGCGCGCGAGAGCTTCGCTTCAGAGAGACACCCCGCGCGCGCGCGCGAG 1428
Db 2152 CCAACAGCGCGCGCGCGAGAGAGCTTCAGGTTTGGGGAAGAGACAACAACCTCCCTCAG 2211
QY 1429 GAGCCCAAGGACCGGAGCGCCCTAGCGGAGCGCGCTGACCGCGCGCGCGCGCGCGCTG 1488
Db 2212 AAGCAGGCGCGATAGACAAGCACTGTATCTTTAGCTTCCCTCAGATCACTCTTGGG 2271
QY 1489 AGCGGCGCGCGCTGAGCCAGTAA 1509
Db 2272 AGCGACCGCGCTGACAAATAA 2292
```

```
RESULT 15
US-08-388-353-1
; Sequence 1, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGirollo, Frank S.
; REGISTRATION NUMBER: 31,345
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
```

```
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9709 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-388-353-1
```

```
Query Match 42.6%; Score 642.2; DB 3; Length 9709;
Best Local Similarity 65.6%; Pred. NC. 1.9e-86;
Matches 998; Conservative 0; Mismatches 493; Indels 30; Gaps 3;

QY 1 ATGGGCGCGCGCGCCAGCATCTCGCGCGGAGAGCTGCACAAAGTGGGGAAGATCCGC 60
Db 790 ATGGGTGCGAGAGCGTCGGTATTAAAGCGGGGAGAAATTAGTAAATGGGAAAAATTCGG 849
QY 61 CTGGCGCGCGCGCGCAAGAAAGCACTACATGCTGAAGCACCTGTGTGGCGCAGCCGAG 120
Db 850 ITAAGGCCAGGGGAAAGAAACAAIATAAACTAAACATATAGTATGCGCAAGCAGGGAG 909
QY 121 CTGAGGGCTTCGCCCTGAACCGCGGCTGCTGGAGACCGCGGAGGCTGCAAGCAGATC 180
Db 910 CTAGAAGATTTCGAGTTAATCCTGGCCTTTTAGAGACATCAGAAGGCTGTAGACAATA 969
QY 161 ATGAAGCAGCTGCGAGCGCGCTGCGAGCCGCGCAGGAGCGTGGCGAGCGCTGTACAAC 240
Db 970 CTGGACAGCTACAACCATCCCTTCAGACAGGATCAGAACAACITAGATCATTTATAT 1029
QY 241 ACCGTGGCCACCTGTACTCGGTGCGAGCGCGGATCGAGTCCGCGACACCAAGAGGCGC 300
Db 1030 ACAATAGCAGTCTCTATTGTGTGATCAAGAGTAGATGTAAGAGACACCAAGGAAGCC 1089
QY 301 CTGCACAAGATCGAGGAGGAGAGCAAGTCCCGACAGAGACCGCAGCAGG----- 353
Db 1090 ITAGATAGATAGAGGAGAGCAAAAGTAAGAAAGGCAAGCAAGCAGGAGCAGCT 1149
QY 354 --CAAGSAGGCGCAGCGCAAGGTGAGCGAGCACTAGCCCATCTGTCAGAACCTCGCAGGC 411
Db 1150 GACACAGAAACACAGCCAGGTGACCGCAAAATACCTCTATGTGCGAAGACCTCCAGGG 1209
QY 412 CAGATGTGCACAGCGCATCAGCGCCCGCGCACCTGAAACCCCTGGGTGAAGTGTAGAG 471
Db 1210 CAAATGTATCATGCGCATATCACCTAGAACCTTTAAATGCATGCGTAAAGTAGTAGAA 1269
QY 472 GAGAAGSCCTTCAGCGCGCGAGGTGA*CCCAICGTTCAAGCGCCCTGAGCGAGGCGCAC 531
Db 1270 GAGAAGCCTTCAGCGCGCGAGGTAAT*ACCCATGTTTTCAGCATATCAGAGGAGCGCAC 1329
QY 532 CCCAGSAGCCTGAACAGCATGTTTCAACACCGTGGCGCGCCACCGCGCCATGCGAGATG 591
Db 1330 CCACAAGATTAAATACCATGCTAAACACAGTGGGGGACATCAAGCAGCCATGCAATG 1389
QY 592 CTGAAGACACCATCAAGCAGGAGGCGCGGAGTGGAGCGCGCTGCACCGCTGCGAGGCG 651
Db 1390 TTAAGAGAGACCATCAATGAGGAAGCTGCAAGATAGATAGATGTCATCCAGTGCATGCA 1449
QY 652 GGCGCGCTGGCGCGCGCAGATGCGCGACCGCGCGGAGCGAGCATCGCGCGCGCAC 711
Db 1450 GGCGCTATTGCGCCAGCGCGAGTGAAGAACCAAGGGGAGTGACATAGCAGGAATCT 1509
QY 712 AGCACCGCTGCAGGAGCAGATCGCGCTGGATGACCCAGAACCGCCCGCTGCGCGCGCAC 771
Db 1510 AGTACCCCTCAGGACACAAATAGGATGGATGCACACATATCCACCTATCCAGTAGAGAA 1569
QY 772 ATCTAAGCGGTGGATCATCTCGGCGCTGAACAAGATCGTGGGATGTACAGCCCGGTG 831
Db 1570 ATCTATAAAGATGGATTAATCTCGGATTAAATATAAATAGTAAGATGTATAGCCCTACC 1629
QY 832 AGCATCTGGACATCGCGCAGGCGCGCAGGAGCGCTTCGGGAGTACGTGACCGCTTC 891
Db 1630 AGCATTTCTGACATAAGACAAAGGAGCCCAAGAACCCCTTTAGAGACTATGTAGACCGATT 1689
```

```
Qy      892  TTCAGAGCCCTGGCGCCGAGCAGCCACCCAGAGCTGAGAGACTGGAGTACCGGAGACC 951
Db      1690  TATAAAGACTGTAGAGCCGAGCAAGCTTCACAAGAGCTAATAAATTTGGATGACAGAAACC 1749
Qy      952  CTGCTGGTGAGAGAACCCAAACCCGACTGCAGACCATCTCTGGCGCTCTGGGCTCCCGG 1011
Db      1750  TTGTTGGTCCAAANTGCGAACCCAGATTGTAGACTATTTTAAAGCATTTGGATCAGAA 1809
Qy     1012  GGCACCTGAGAGAGATGATGACCGCTGCCGGG/GTGGGGGGCCCGCCGACAGAGGT 1071
Db      1810  GGCACACTAGAGAAATGATGACAGCACTGTCAGGAGTGGGGAGCCCGCCGATAAAGCA 1869
Qy     1072  CGGCTGCTGCCGAGGATGAGCCAGG---CTAACAGGCTGAACATGATGATGAGAG 1129
Db      1870  AGAGTTTGGCTGAGGCAATGAGCCAGTAAACAATGCAAGCTACCATTAATGATAGAAA 1929
Qy     1129  AGCAACTTCAAGGGCCCGCGCCCAAGTCAAGTGGTCAACTGGGCAAGGAGAGTCCAC 1188
Db      1930  GGCATTTTAGGAACCAAGAAAGAGACTGTTAAGTGTTCATTTGCGAAGAGAGAGCCAC 1989
Qy     1189  ATGCCAAGAACTGCGCGCCCGCCCGGAGAGAGAGGGCTGCTGGAAGTGGCCAGAGAGGG 1248
Db      1990  ATAGCCAAAATTCAGGGGCCCTAGCAAAAAGGGCTGTGAAATGTGSAAGSAAGCA 2049
Qy     1249  CACGAGATGAGGACTGACCGAGCCGAGCCGAGCCCAACTTCCTGGGCAAGATCTGGTCCAG 1308
Db      2050  CACCAANTGAAGATGTGCTGAGACAGACGGCTAATTTTTAGGGAAGATCTGGGCTTCC 2119
Qy     1309  CACAAGGGCCCGCGCCCAACTTCCTGCAAGAACCGCAGGAGCCCGCCCGCCGACCGT 1368
Db      2110  CACAAGGGAAGGCAGGGAATTTCTTCAGAGCAG-----AGCAGAG 2151
Qy     1369  CCCACGCCCGCCCGCCGAGAGCTTCGCTTCGAGGAGACACCGCCCGCCCAAGCAG 1428
Db      2152  CCACAGCCCGCCCGAGAGAGAGCTTCAGGTTTGGGGAAGAGACACAACTCCCTCTCAG 2211
Qy     1429  GAGCCCAAGGACCGGAGCCCTACCGCGAGCCCTGACCGCCCTGGCGAGCCTGTTCCGC 1488
Db      2212  AAGCAGAGCCGATAGACAGAGAACTGTATCCTTTAGCTTCCTCAGATCACCTTTTGGC 2271
Qy     1489  AGCGGCCCTGTAGCCAGTAA 1509
Db      2272  AGCGACCCCTCGTCACAATAA 2292
```

Search completed: September 22, 2003, 22:53:53  
Job time : 109.469 secs





FEATURES		Location/Qualifiers		KEYWORDS		SOURCE	
source		1..933		Gasterosteus aculeatus (three spined stickleback)		EST.	
		/organism="Zea mays"		Gasterosteus aculeatus		Fukuyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		/mol_type="genomic DNA"		Fukuyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;	
		/strain="B73"		Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;	
		/db_xref="taxon:4577"		Gasterosteidae; Gasterosteus			
		/clone="2MMBT3458K20"		1 (bases 1 to 1219)			
		/clone_lib="ZM_0.6_1.0_KB"		Kingsley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickson, M.,			
		/note="Vector: pCR4-TOP0; Site_1: EcoRI; 0.6-1.0 kb high;		Schmutz, J. and Myers, R.M.			
BASE COUNT		Cof selected genomic DNA library"		Expressed sequence tags from Gasterosteus aculeatus			
ORIGIN		198 a 276 c 309 g 150 t		Unpublished			
				Contact: Kingsley, DM			
				HHMI and Department of Developmental Biology			
				Stanford University School of Medicine			
				Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA			
				Tel: 650 725 5954			
				Fax: 650 725 7739			
				Email: kingsley@cngm.stanford.edu			
				Plate: 43			
				High quality sequence start: 4			
				High quality sequence stop: 792.			
FEATURES		Location/Qualifiers					
source		1..1219		/organism="Gasterosteus aculeatus"			
				/mol_type="mRNA"			
				/strain="Salinas river, CA"			
				/db_xref="taxon:69293"			
				/clone="CD443-C08"			
				/sex="mixed male and female"			
				/tissue_type="heads and internal organs combined"			
				/dev_stage="adult"			
				/clone_lib="SHGC-CDA"			
				/note="Vector: lambda ZAP Express/pBK-CMV; Site_1: EcoRI			
				(5' adaptor); Site_2: XhoI (3' linker primer); The mixed			
				organ cDNA library was generated using the ZAP-cDNA method			
				by Stratagene. First strand cDNA synthesis was primed with			
				a 50 bp linker primer containing an oligo dT sequence			
				preceded by a synthetic XhoI site. 5 prime adaptors were			
				used containing an EcoRI cohesive end. The finished cDNAs			
				were inserted in to the ZAP express vector			
				unidirectionally in the sense orientation with respect to			
				the lacZ promoter of pBK-CMV. An amplified library was			
				prepared from approximately 3 million primary clones in			
				the lambda ZAP Express vector. In vivo excision was then			
				used to generate individual pBK-CMV phagemid clones for			
				EST sequencing."			
BASE COUNT		201 a 440 c 390 g 188 t					
ORIGIN							





```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruher, C., Jessop, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
RP 191 9-006 EVRY cedex - France
Email: seqret@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS001015CA02NP1.
Unpublished

FEATURES
source
LOCATION/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS001015B03"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 116 a 88 c 93 g 398 t 506 others
ORIGIN
Query Match 5.3%; Score 79.8; DB 13; Length 1201;
Best Local Similarity 9.6%; Pred. No. 0.00071;
Matches 59; Conservative 303; Mismatches 234; Indels 0; Gaps 0;

Qy 894 CAAGACCTCGCCGCCAGGACGCCACCCAGGAGCTGAAGACTGGATGACGAGATCC 953
Db 1201 VAAVSSAVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1142
Qy 954 GCTGTCGACAGCCACCCGACTCGACAGACTCTCGCGCTCGCGCCGCCGCC 1013
Db 1141 SVVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1082
Qy 1014 CACCTCGAGGAGATGATACCGCTCCAGGCGTGGCGGCCGCCGCCACAGGCGG 1073
Db 1081 ASVSSSSVAAVAVAAVAVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1022
Qy 1074 CGTCTGGCGAGCGATGAGCCAGGCCACAGCGTGAACATCATGATGAGAGACAA 1133
Db 1021 SSSANSSVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 962
Qy 1134 CTTCAGGGCCCGCGCGACGCTCAAGTCTCACTCGGCAAGGAGGCCACATCC 1193
Db 961 MSSAVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 902
Qy 1194 CAAGAAGTGGCGGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGAAGGAGGCCACCA 1253
Db 901 SAAAVASAVSAAAVVSSVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 842
Qy 1254 GATGAAGACTGACAGCGAGCCAGGCCAAGCTCTCTGGGCAAGATGTGGCCCGCCACAA 1313
Db 841 SSVAVSAAAVSSVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 762
Qy 1314 GGGCCCGCCGGCACTTCTGCAAGACCGCAGCGAGCCCGCCCGCCCGCCCGCCAC 1373
Db 761 VSSVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 722
Qy 1374 CGCCCGCCCGCCCGAGAGCTTCCTCTCGAGGAGACACCCCGCCCGCCCGCCCGAGGAGCC 1433
Db 721 SAAVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 662
Qy 1434 CAAGACCGGAGCCCTACCGCGAGCCCTCAGCGCCCTCGGCAAGCTGTTCGCGACGG 1493

```

```

Db 661 SASSSAAVAAANSANNNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 602
Qy 1494 CCCCTGAGCCAGTAA 1509
Db 601 NSSNSNANNNNANAAA 586

RESULT: 6
AY106831
LOCUS AY106831
DEFINITION Zea mays P00142079 mRNA sequence.
ACCESSION AY106831
VERSION AY106831.1 GI:21209909
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2299)
AUTHORS Hainey, C.F., Solau, M., Miao, S.H., Vogel, J.M., Whittitt, M.S.,
Arthur, L.W., Hanarey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
REFERENCE 2 (bases 1 to 2299)
AUTHORS Coe, E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org: ZmDB, www.zmdb.iastate.edu; FIGR,
www.tlgr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES
source
LOCATION/Qualifiers
1..2299
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:638374"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project."
BASE COUNT 475 a 790 c 658 g 376 t
ORIGIN
Query Match 5.3%; Score 79.8; DB 11; Length 2299;
Best Local Similarity 43.9%; Pred. No. 0.00077;
Matches 495; Conservative 0; Mismatches 617; Indels 15; Gaps 3;

Qy 122 TGAGAGGCTTCGCCCTGAACCCCGGCTCTGTGAGACGCCGAGGCTGCAAGCAGATCA 181
Db 767 TGAGGCCCGGCCCATCATGGAGCATCTCTGATGGCAGCTCTTCATGAAGCAGGCCA 826
Qy 182 TGAAGCAGCTGCAGCCCGCCCTGCACCGCGGACCGGAGGCTGCCAGCCCTGTACAACA 241
Db 827 AGGAGGTGAACGCCCATGGACCGGCTCTGAAAGCCGAGGAGGACAGGTACGGCTCCGCA 886
Qy 242 CCGTGCCACCTGTACTGCTGCACGCCGGCATCGAGGTTCGCGACACCAAGAGGCC 361
Db 887 CGTCGCCAGTGGCTCGGCCCGCCAGATCGAGGTATCTCCGCCCGCCCAAGTCCATCG 946
Qy 302 TGSACAAAGATCGAGGAGGAGCAGAACAGTCCAGCAGAGAGACCCAGCAGCAGGAGG 361

```

```

Db 947 AGCAGAGTCAACTCGGTCACACACACCCCTGTCATCGACGTCACACCGTGGCAAGCCGC 1006
Qy 362 CGAGAGCGAAGTGTAGCGAGAACCTACCTCATCTGCGAAGACCTCGAGGCCGAGATGGTTC 421
Db 1007 TGCACGCGGCAACTTCCAGGGCACCCGATCGCGGTCTCCATGACACAGCGCGCGCTGG 1066
Qy 422 ACCAGGCGATCAGCGCGGACCTGAAAGCCGCTGGTGAAGTGAACGAGAGAGAGCGCT 481
Db 1067 CGCTCGGCACATCGGCAAGCTCATGTTCGCCAGTTCCTCGAGGTGGTCAACGAGTCT 1228
Qy 482 TCAGCGCGGAGGTGATCGCCCATGTTACGCCCTGAGCGAGGCGCCACCCCGCCAGAGC 541
Db 1227 ACAACACGGGCTCACTCCAACTGCGGCGGACGCCAACCGGACCGGCTGGACTAGGG 1186
Qy 542 TGAACAGGATGTTGACACCGTGGCGGCGACACAGCGGCCCATCGAGATGCTGAAGGACA 601
Db 1187 TCAAGGCGACCGATCGGCGATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1246
Qy 602 CCATCA-----ACGAGGAGGCGCGCGAGTGGGACCGCGCTGCAACCGCTGCAAGCGCGGC 655
Db 1247 CCATCAACACCGTCCAGAGCGCGGAGGACGACACACAGGACGTGAACCTCCCTCGGC 1306
Qy 656 CGTGCGCGCGCGGAGATCGCGGACCGCGCGGAGCGGACGACGCGCGCGGCGCACACGCA 715
Db 1307 TCGTCTCGCGCAGGAGACCGCGGCGGTGGACATCCTCAAGCTCATGTGCTCCAGCT 1366
Qy 716 CCGTGCAGGAGGATCGCGCTGGATGACGACGACGACCGCGCGCTGCGCGCGGCGATCT 775
Db 1367 ACATGTCGCGCTGTGCCAGCGCGTGACCTCGCGCACCTCGAGGAGAACCTCAAGAGCG 1426
Qy 776 ACAAGCGGTGGATCATCTTGGGCTGAACAAGATCTGTGGGATGTACAGCCCGCTGAGCA 835
Db 1427 CCGTCAAGAGCTCGTGTATGGCGTGGCCAGGAGGTGTGACACCAACCGCTCGCGCGCG 1486
Qy 836 TCGTGGACATCGCGAGGCGCGCAAGGAGCGCTTCCGCGACTACGTGACCGGCTTCTCA 895
Db 1487 ACCTCCACGCGCGCTTACGAGAGAGCGCTTCCGCGACTACGCGCGAGGCGCG 1546
Qy 996 AGACCCCTCGCGCGGAGCAGCGCCACCGAGGACGTGAAGACCTGATGATGACCGGAGA----- 949
Db 1547 TGTACGGGTACTACGAGGACCGCTTGCAGCGCCCACTCGCGCGCGCTGATGAAGAAGATCGGG 1606
Qy 950 CCGTGTGTGCAAGCGCGCAACCGGACTGCAAGACCATCTCGCGCGCTCTCGCGCGCG 1009
Db 1607 CCGTGTGTGAGGACCGCGCTCGCGAGCGCGGAGGCGGAGAGGACGCGCGCTCGCG 1666
Qy 1010 CGCGACCTGAGGAGATGATGACCGCTTGCAGGCGCGTGGCGCGCGCGCGCGCAAGG 1069
Db 1667 TGTCTCCAGATCAACAGGTCGAGGAGGAGCTGCGCGAGGCGCTGCGCGAGGATCG 1726
Qy 1070 CCGCGGTGTGCGGAGGAGATGAGCGAGCGCAACAGGCTGATGATGATGATGATGAT 1129
Db 1727 AGCGCGCGCGGTGGCTTTCAGACCGCGCGCGCGCGCGATCGCAACAGATCAAGGAGA 1786
Qy 1130 GCAACTTCAAGGCGCGCGCGCGCAAGCTCAAGTTCGTAACCTCGCGAGGAGGCGCA 1169
Db 1787 GCGGTGTA---CCGCTGTACCGCTTATCGCGAGGACCTCGCGCGCGCTGTACCTGA 1643
Qy 1190 TCGCCAAAGTCCGCGCGCGCGCGCGCAAGAGGCGCTCTGGAAGTGC 1236
Db 1844 CCGCGGAGAGCTCAAGTCCCGCGCGGAGGAGTGCACAGAGGTGTT 1890

```

## RESULT 7

BF259495

LOCUS

DEFINITION BF259495 759 bp mRNA linear EST 22-OCT-2001  
 HVSMEF0019H19f Hordeum vulgare seedling root EST library HVSNA0007  
 (Etiolated and unstressed) Hordeum vulgare subsp. vulgare cDNA  
 clone HVSMEF0019H19f, mRNA sequence.

ACCESSION

BF259495

VERSION

BF259495.2

GI:13120022

KEYWORDS

EST.

SOURCE

Hordeum vulgare subsp. vulgare

## ORGANISM

Hordeum vulgare subsp. vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
 ; Triticeae; Hordeum.

## REFERENCE

1 (bases 1 to 759)

## AUTHORS

Wing, R., Close, T.J., Kleinholz, A., Wise, R., Begum, D., Frisch, D., Yu  
 Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton  
 R.D., Oates, R. and Main, D.

Development of a genetically and physically anchored EST resource  
 for barley genomics: Morex unstressed seedling root cDNA library

## JOURNAL

Unpublished

On Nov 16, 2000 this sequence version replaced g11188603.

## COMMENT

Contact: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7283

Fax: 864 656 4293

Email: rwing@clemson.edu

Total hg bases = 388

Seq primer: AATTAACTCTCACTARAGGG

High quality sequence stop: 758.

## FEATURES

source

1..759  
 Location/Qualifiers  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="mRNA"  
 /cultivar="Morex"  
 /db\_xref="taxon:112509"  
 /clone="HVSMEF0019H19f"  
 /lab\_host="TJC121"  
 /tissue\_type="Seedling root"  
 /clone\_lib="Hordeum vulgare seedling root EST library  
 HVCNA0007 (Etiolated and unstressed)"  
 /note="vector: lambdaZAP; Site.1: EcoRI; Site.2: XhoI;  
 Seeds were surface sterilized then germinated under axenic  
 conditions in the dark at room temperature on filter paper  
 with water, tystatin and cefotaxime in covered  
 crystallization dishes. Five-day old seedling roots were  
 then harvested. total RNA was prepared, poly(A) RNA was  
 purified, one primary unamplified cDNA library was made,  
 and 1 million pfu were in vivo excised to give pBluescript  
 SK(-) cDNA phagemids. These steps were performed in the TJ  
 Close laboratory at the University of California,  
 Riverside (Choi, Close, Fenton). Phagemids were plated and  
 picked at the Clemson University Genomics Institute (CUGI)  
 (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA  
 preparations, DNA sequencing and sequence analysis were  
 performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates  
 Rambo, Main). The sequence has been trimmed to remove  
 vector sequence and contains a minimum of 100 bases of  
 phred value 20 or above. For more details on library  
 preparation and sequence analysis see  
 http://www.genome.clemson.edu/projects/barley. To order  
 this clone see http://www.genome.clemson.edu/orders Also  
 see Close TJ, Wing R, Kleinholz A, Wise R (2001)  
 Genetically and physically anchored EST resources for  
 barley genomics. Barley Genetics Newsletter 31:29-30.  
 (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html) \*

BASE COUNT

ORIGIN

158 a 282 c 184 g 133 t 2 others  
 Query Match 5.18; Score 76.4; DB 10; Length 759;  
 Best local similarity 47.8%; Pred. No. 0.0025;  
 Matches 283; Conservative 0; Mismatches 303; Indels 6; Gaps 2;

Qy 124 GAGGCTTCGGCTGAACCGCGGCTCTGAGAGACCGCGGAGGCTCAAGAGATCATG 183

Db 13 GAGATCACCGCCACACCCATGAACTGATGACGAGCTGCGCGGAGGCGGCGAGTTC 72

Qy 184 AAGCAGCTGCAGCGCGGCTGACAGCGGACCGGAGGAGCTGCGGACGCTGTACAACAC 243

Db 73 CTCACATGTGCTCTCAAGCTCATCGGCGCCCAAGAAGACCATGGAGATCGGCGCTACAG 132



Db	931	CCGCTGAGGTGCTCAAGTTCGCCATCTGCTCAAGCGGCTCTCCGCGCAAGGCTGTGAACAA	943
QY	1245	GGGCGCCACGATGAAGACATGGCAACCGCCAGCGCCAGCAACTTCCTGGGCGAAGATCTGGCC	1304
Db	941	CAATCAACAACAAGACGGCTTCACCAACCGCACCGCACTACGGCAAGGCGGAG-----	942
QY	1305	CAGCGACAAGGGCGGCGCCCGCGCAACTTCGTGCAAGACCGCAACGAGCGCGCGCGCGCAAC	1364
Db	993	----CGAGAGCGGAGTGGCGCCACGGCAGACAGAGAGCGGTGCACGCGCTCAACACGAGCGAC	1348
QY	1365	CGTCCCGACGCGCCCGCCCGCCGACGAGCTTCGCTTCGAGGAGGAGACACCGCCCGCCCGCA	1424
Db	1049	CGCCACCTCCGACCTCTTCGCGCAACCAACGAGGCTACCGGAGCTGTGGAGCTCGCGGA	1105
QY	1425	CGAGGACCCCAAGGACCGCGAG	1446
Db	1109	CGAGCGGCCAAGCGCGCGAG	1130
RESULT 9			
LOCUS	BG314500		
DEFINITION	WHE2495_E11_121ZS Triticum monococcum cDNA clone WHE2495_E11_121, mRNA linear EST 16-APR-2001		
ACCESSION	BG314500		
VERSION	BG314500		
KEYWORDS	Triticum monococcum		
SOURCE	EST		
ORGANISM	Triticum monococcum		
REFERENCE	Anderson,O.D., Chao,S., Dubcovsky,J., Echenique,J., Han,P.S., Hsia		
AUTHORS	,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausen,C.J., Seaton,C.L.,		
TITLE	The structure and function of the expressed portion of the wheat		
JOURNAL	genomes - Early reproductive apex cDNA library from Triticum		
COMMENT	monococcum		
	Unpublished		
	Contact: Clin Anderson		
	US Department of Agriculture, Agriculture Research Service, Pacific		
	West Area, Western Regional Research Center		
	800 Buchanan Street, Albany, CA 94710, USA		
	Tel: 5105595773		
	Fax: 5105595818		
	Email: andersn@pw.usda.gov		
	Sequence have been trimmed to remove vector sequence and low		
	quality sequence with phred score less than 20		
	Seq primer: Stratagene SK primer.		
	location/Qualifiers		
	1..509		
FEATURES	/organism="Triticum monococcum"		
source	/mol_type="mRNA"		
	/cultivar="DV92"		
	/db_xref="taxon:4568"		
	/clone="WHE2495_E11_f21"		
	/tissue_type="Early reproductive apex"		
	/dev_stage="Seven week-old plants"		
	/lab_host="E. coli XLOLR"		
	/clone.lib="Triticum monococcum early reproductive apex		
	cDNA library"		
	/note="Vector: Lambda Uni-ZAP XK, excised plasmid:		
	Site_1: EcoRI; Site_2: XhoI; The tissue, total RNA, and		
	poly(A) RNA were prepared from apex at double-ridge stage		
	to terminal-spikelet stage during transition from		
	vegetative state to flower state, a cDNA library was made,		
	and the cDNA clones were in vivo excised at the		
	University of California, Davis (V. Echenique, B. Stamova		
	, J. Dubcovsky). Plasmid DNA preparations and DNA		
	sequencing were performed in the OD Anderson lab (all		
	other authors)."		





1057 CTTTCATGAAGCAGGCGCAAGAAATTTCAACGAGCTGGAACCCGCTGCTGAAGCCCAAGCAGGA 1116  
QY 519 CGAGGGCCCAACCCCGAGACCTTGACACACATGTTGA---ACACGCTGGGGGCGGACCA 575  
Db 1117 CAGGTACCGCTGCGGACGTCGGCGCATGGCTGGCGCCCGCAGATCGAGGTCTATGDCGCG 1176  
QY 576 GCGCCCATGAGATGCTGAAGGACACATCAAGAGGAGCGCCGCGAGTGGAGCGCCCT 635  
Db 1177 GCCCACCAAGTCCATCGAGCGGAGTCAATCGCTGAAGACACACCGGTCTCATGAGCT 1236  
QY 636 GCACCCCTGAGCGCGGCGCGCTGGCGCCCGGCGCAATGCGGACCGCGCGCG ---CGAG 692  
Db 1237 CCACCGCGCAAGCGCTGCAGCGCGCACTTCAGGGACCGCCCATCGGCGGTGCTCAT 1296  
QY 693 GCACATCGCGCGCGCACACACCGCTCCAGAGCAGATCCCTGATGACCAACACCC 752  
Db 1297 GGACAAACCGCGCTCGCCCATCGGCAATCGGAAGCTCATGTCGCGGAGTTCTTCGA 1356  
QY 753 CCGCCCTGCGCGTGGCGACATCTCAAGCGGTGATCATCTGGGCTTGACAAAGATCGG 812  
Db 1357 GCTCTCAACGAGTTCTACAAACAGGSGCTCAGCTCCACCTGCGCGGACCGCAACCC 1416  
QY 813 GCGGATGTACAGCCCGGTGAGATCTCTGGACATCGGCCAGCGCCCGAAGAGCGCTTCGG 872  
Db 1417 CAGCCTGGACTAGCGCTTCAAGGGCAGCGAGATCGCCATGGCTCTACTGCTCCGAGCT 1476  
QY 873 CGACTACGTGGACCGCTTCTTCAACACCGCTCGCGCGCAGCAGCGCCACCGAGAGCTGAA 932  
Db 1477 CCAGTACCTGG---GCACCCCATCACCAACACGTCGAGAGCGGGAGCAGCACACCA 1533  
QY 933 GAACGTGATGACCGAGACGCTGCTGGTGCAGAGCGCCCAACCGCACTGCAAGACCATCT 992  
Db 1534 GGAGTGAATCTCCCTGGGCTGCTGTCGCCAGGAGACCGCGCGAGGCGATCGACATCT 1593  
QY 993 GCGGCTCTCGCGCGCGCGCACCTCGAGGAGATGATGACCGCTGCCAGGCGCTGG 1052  
Db 1594 GAAGCTAIGTCTCACCTATCATCTGGGCTGTGGCAGCGGTGGACCTGCGCCACCT 1653  
QY 1053 CGCGCCCGCGCACAAAGCGCGCTGCTGGCGAGCGATGACCGAGGCGCAACAGCGTGA 1112  
Db 1654 CGAGGAGACATCAAGGCTGCTGTAAGAACACCGTGACCCAGGTGGCCAAAGAGTGGT 1713  
QY 1113 CATCATGATCGAGAGACGACTCAAGGCGCGCGCGCGCAAGCTCAAGTCTTCAACTG 1172  
Db 1714 GACCATCAACCGCTCGGCGAGCTCTCCAGCGCGGCTTCAGCGAGAGAGAGCTGATCA 1773  
QY 1173 CGCGAAGGAGGCGCACATCGG-----CAAGAACTCGCGCGCGCGCG 1214  
Db 1774 CGCCATCGACCGCGAGGCGGCTTCAAGTACGCGAGGAGCGCGCGCGGCGAGCTTCC 1833  
QY 1215 CAAGAAGGCTCTCTGAGTCTGGGCAAGGAGGCGCGCAGATGAGGAGCTGACCGAGCG 1274  
Db 1834 GTGATGACAGAGTCTGCGCGCTGCTGTGGACACCGCTTCAGCAGCGCGAGCGGA 1893  
QY 1275 CGAGGCAACTTCTTGGGCAAGATCTGCGCGAGCGCAAGGCGCGCGCGCGCAATTCCT 1354  
Db 1894 CGGGAGGCGCTCGGTCTCTCAAGATCACAGTCTCGAGAGAGAGCTCGCGCGGTCT 1953  
QY 1335 GCAGAACCGAGCG 1394  
Db 1954 GCGCAGGAGGTGGAGCG 2013  
QY 1395 CGCTTCGAGAGACACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1454  
Db 2014 CGGATTCGGGAGAGCGGCTGCTTCCGCTGTACCGCTTCTGCGGAGGAGCTCGGCTG 2073  
QY 1455 CGAGCGCTGACCGCGCTGCGCAGCTGTTCGGAGCG 1492  
Db 2074 CGTGTTCGAGCGGAGAGCTCAAGTCCCGCGCG 2111

RESULT 12  
CB870888

LOCUS CB870888 649 bp mRNA linear EST 23-APR-2003  
DEFINITION HC15104w HC Hordeum vulgare subsp. vulgare cDNA clone HC15J04  
5-PRIME, mRNA sequence.  
ACCESSION CB870888  
VERSION CB870888.1 GI:30072868  
KEYWORDS EST.  
SOURCE Hordeum vulgare subsp. vulgare  
ORGANISM Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
REFERENCE 1 (bases 1 to 649)  
AUTHORS Varshney, R.K., Zhang, H., Burton, R., Stein, N., Landridge, P. and  
Graner, A.  
TITLE Parley ESTs from coleoptile tissue  
JOURNAL Unpublished  
COMMENT Contact: Stein, Nils  
Molecular Markers Group, Department Genbank  
Institute of Plant Genetics and Crop Plant Research (IPK)  
Cottbusstr. 3, 06466, Gatersleben, Germany  
Tel: 039482-5522  
Fax: 039482-5595  
Email: stein@ipk-gatersleben.de  
Insert Length: 649 Std Error: 0.00  
Plate: 15 row: 3 column: 4  
Seq primer: T7.  
FEATURES  
Location/Qualifiers  
1..649  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/strain="cv tokak"  
/db\_xref="GABI:549289"  
/db\_xref="taxon:112509"  
/clone="HC15J04"  
/tissue\_type="Root"  
/dev\_stage="3 week old"  
/clone\_lib="HC"  
/note="6 and 10 hour drought stress by placing plants on  
moist paper (75% rel. humidity) in light"  
BASE COUNT 128 a 253 c 175 g 92 t 1 others  
ORIGIN

Query Match 4.9%; Score 74.2; DB 14; Length 649;  
Best Local Similarity 47.6%; Pred. No. 0.0098;  
Matches 262; Conservative 0; Mismatches 304; Indels 6; Gaps 2;  
QY 124 GAGGGCTCGCCCTGAACCCCGCGCTGTGTGAGACCGCGGAGGCTGCAAGCAGATCAIG 183  
Db 45 GAGATCAACCGCAACCAACCCATGGAACCTGATGACGACGCTCGCGGACGAGGCGCAGTIC 104  
QY 184 AAGCAGCTGAGCGCGCGCTGTCAGACCGGACCGGAGGAGCTCGGAGGCTGTACACAGC 243  
Db 105 CTCACATGCTGTCAAGCTCATCGCGCGCAAGAACATGAGATCGGCGCTCTACAGC 164  
QY 244 GTGGCCACCTCTACTGCTGCTGACCGCGCGCATCTGAGGTCGCGGACAGGAGGCGCGT 303  
Db 165 GGTACTCTGCTCGCCACCGCGCTCGCCATCCCGGAGCAGCGCACCATCTTGGCCATG 224  
QY 304 GACAAGATCGAGGAGGAGCAGAACAGTCCCAGCAGAAAGCCCGCAGCGCCCAAGAGGCC 363  
Db 225 GACATCAACCGCGAGAACTACGAGCTGGGGCTGGCGTGCATCGAGAAAGCGCGCGTGGCG 284  
QY 364 GACGGCAAGTGAAGCAGAACTACCCATCGTCGAGAACTCGAGGCGCGAGATGCTGAC 425  
Db 285 CACAAGATGACTTCGCGAGGCGCGCGCGCTCCCGCTCTCTGAGCGCGCTCTCTGAGGAC 344  
QY 424 CAGCGCATAGCGCGCGCGCGCTGAAAGCTGGGTGAAGGTGATGAGGAGAGGCGCTTC 483  
Db 345 GAGCGCAACCGCGCACCTTCGACTTCGCTTCGCTGGAGCGCGCGAGCAGACATACCTC 404  
QY 484 AGCCCGGAGTGATCCCCATGTTTCACCGCGCGCTGAGCGAGGCGCGCTACCGCGAGACCTG 543  
Db 405 AACTACCGAGGCGCTCATGAAGTCTGTCAGGTCGGGCGCGCTCTCTGCGGTACGAC--- 461

```

544 AACACGATGTGTGAACACCTGGCGGCCACCGCGGCATCGACATGCTGAAGACACC 603
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
462 AACACCTCTGGAACGCTCGCTGCTCGCGCGGACGCGCCCATCGCAAGTACATC 521
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
504 ATACAGGAGGAGCGCGGAGTGGGACCGCTGCACCCCGTGCAGGCGCGC---CCGCTG 560
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
522 CGTACTACCGCGACCTCGCTCGACCTCAACAGGCGCTCGCGGACGACGCGGCTC 581
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
661 GCCCGCCGCGACATGCGCGACCGCGCGGACGACATCGCGCGCGCACCA 712
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
582 GAGATCTGCCAGCTCCCGCTCGCGAGCGCATACCTCTCGCGCGACCCA 513
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
BI960118 537 bp mRNA linear EST 22-OCT-2001
LOCUS HVSME0023E18f Hordeum vulgare rachis EST library HVCNMA0015
DEFINITION (normal) Hordeum vulgare subsp. vulgare cDNA clone HVSME0023E18f,
mRNA sequence.
ACCESSION BI960118 GI:16311373
VERSION R1960118
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
1 (bases 1 to 637)
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Chin, A., Begum, B.,
Frisch, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, J., Simmons
J., Oates, K. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex rachis cDNA library
Unpublished
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
106 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total nt bases = 579
Seq primer: AATTAACCTCTCAATAAGGG
High quality sequence start: 5
High quality sequence stop: 633.
FEATURES
Location/Qualifiers
1..637
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/db_xref="taxon:12509"
/clone="HVSME0023E18f"
/tissue_type="Rachis"
/lab_host="TJ121"
/collection="Hordeum vulgare rachis EST library HVCNMA0015
(normal)"
/notes="Vector: pBluescript SK(-); Site 1: FcoRI; Site 2:
XhoI; Plants were grown at Washington State University,
Pullman, WA in a greenhouse, the rachises were excised and
frozen in liquid nitrogen (Kleinhofs lab). In the TJ Close
lab at the University of California, Riverside total RNA
was prepared, poly(A) was purified, one primary
unamplified cDNA library was made, and 1 million pfu were
in vivo excised to give pBluescript SK(-) cDNA phagemids
(Chin). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Frisch, Henry, Simmons, Rambo, Main). The
sequence has been trimmed to remove vector sequence and
contains a minimum of 100 bases of phased value 20 or
above. For more details on library preparation and

```

```

sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
Barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)*
BASE COUNT 125 a 249 c 158 g 105 t
ORIGIN
Query Match 4.9%; Score 74; DB 12; Length 637;
Best Local Similarity 48.9%; Pred. No. 0.0042;
Matches 259; Conservative 0; Mismatches 265; Indels 6; Gaps 2;
QY 186 GCAGCTGCACCCGCCCTGCAGACCGGACCGAGGAGTGGCGAGCTGTATTAACACCGT 245
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13 GCATGCTGCTCAAGCTCATCGGCCCAAGACCAIGGAGATCGGCGTCTACACGG 72
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 246 GGCACCCCTGTACTGGTGCACCGCGCATCGAGTCCCGGACACCAAGGAGGCCCTGGA 305
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 73 CTACTCCCTGCTCGCCACCGCGCTCCCGACCGACGACCATCTTGGCCATGGA 132
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 306 CAAGATCGAGGAGGAGAGAACAAAGTCCCGAGAGAACCCAGACGCGCAAGGAGCGGA 365
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 133 CATCAACCGCGAGAACTACGAGCTGAGCGTGGCGTGCATCGAGAGCGCGGTGGCGCA 192
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 366 CGGCAAGGTGAGCGAAGTACCCCATCTGTCGAGAACTGCGAGGCGCAGATGGTGCACCA 425
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 193 CAAGATCGACTTCGCGGAGGCGCGCGCTCCCGCTCTGGACGCGCCCTCTCGAGAGGA 252
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 426 GGCATCAGCGCCCGCACCCCTGAACCCCTGGGTGAAGTGTATCGAGGAGAGGCCCTTCAG 485
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 253 GSCCAACCGGACCTTCGACTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 312
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 486 CCGCGAGGTGATCCCATGTTTACCGCCCTGAGCGAGGCGCGCCACCCCGGAGACTGAA 545
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 313 CTACCAAGCGGCTCTCAAGCTCTCAAGTCTCGCGCGGCTCTCTCTCTCTCTCTCTCTCTCT 369
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 546 CAGCATGTTCAACACCGTGGCGGCGCACGAGCGCGCATTCAGATGCTCAAGGACACCAT 605
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 370 CACCTCTGGAACGGCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 429
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 506 CAACGAGGAGCGCGCGAGTGGGACCGCTGCACCCCGTGCAGGCGCGCGC---CGTGGC 662
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 430 CTACTACCGGACTTCGTCCTCAAGCTCAAGAGGCGCTCGCGCGGACGACGCGGTCGA 489
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 663 CC CGCGCAGATCGCGGACCCCGCGGCGGACGACATCGCGCGGCGCACCA 712
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 490 GATCTGCCAGCTCCCGCTCGCGGACGGCATCACCTCTCTCGCGCGCGGCA 539
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
BF484304 562 bp mRNA linear EST 06-DEC-2000
LOCUS WHE2321_F12_K232S Wheat pre-anthesis spike cDNA library Triticum
DEFINITION WHE2321_F12_K232S mRNA sequence.
ACCESSION BF484304.1 GI:11567605
VERSION BF484304
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
1 (bases 1 to 562)
Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han
P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,
Seaton, C.L. and Tong, J.C.
The structure and function of the expressed portion of the wheat
genomes - Pre-anthesis spike cDNA library
Unpublished
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific

```

West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA  
Tel.: 5105959773  
Fax: 5105959818  
Email: andersn@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low  
quality sequence with phred score less than 20  
Seq primer: Stratagene SK primer.

# FEATURES

location/Qualifiers

1..562  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="Chinese Spring"  
/db\_xref="taxon:4565"  
/clone="WHE2321\_F12\_K23"  
/tissue\_type="Spike before anthesis"  
/dev\_stage="Adult p-ant"  
/lab\_host="E. coli SOLR"  
/clone\_lib="Wheat pre-anthesis spike cDNA library"  
/note="Vector: Lambda Uni-2AP XR, excised phagemid;  
Site 1: EcoRI; Site 2: XhoI. Plants were grown in the  
greenhouse. Whole spike with awns trimmed. White, green  
and yellow anther were collected and total RNA, and  
poly(A) RNA were prepared, a cDNA library was made, and  
the cDNA clones were in vivo excised to give pBluescript  
phagemids in the TJ Clonase lab (Chol. Clonase, Fenton) at  
the University of California, Riverside. Plasmid DNA  
preparations and DNA sequencing were performed in the OD  
Anderson lab (all other authors)."

108 a 215 c 154 g 85 t

# BASE COUNT

## ORIGIN

Query Match 4.9%; Score 73.8; DB 10; Length 562;  
Best Local Similarity 49.0%; Pred. No. 0.0066;  
Matches 257; Conservative 0; Mismatches 262; Indels 6; Gaps 2;

QY 191 TGCAGCCGCGCTGCAGACGGCCACGAGGAGCTGGCAGCCCTGTACACACCGTGGCCA 250  
DB 191 TGCAGCCGCGCTGCAGACGGCCACGAGGAGCTGGCAGCCCTGTACACACCGTGGCCA 250  
QY 35 TGCTGCTCAAGCTATGCGGCCAAGAACCATCGAGATCGGCTGTACACCGGTACT 94  
DB 35 TGCTGCTCAAGCTATGCGGCCAAGAACCATCGAGATCGGCTGTACACCGGTACT 94  
QY 251 CCCTGTACTGCTGCACCGCGGCTGCGGAGTCCGAGACCAAGGAGCGCTGGAGAAGA 310  
DB 251 CCCTGTACTGCTGCACCGCGGCTGCGGAGTCCGAGACCAAGGAGCGCTGGAGAAGA 310  
QY 95 CCTGCTCGCCACCGCGTCCGATCCCGGACGAGCCATCTTTGGCCATGACATCA 154  
DB 95 CCTGCTCGCCACCGCGTCCGATCCCGGACGAGCCATCTTTGGCCATGACATCA 154  
QY 311 TCGAGGAGGAGCAACAAGTCCACGACGAGAGCCACGAGGCGGAGCGGCGGCA 370  
DB 311 TCGAGGAGGAGCAACAAGTCCACGACGAGAGCCACGAGGCGGAGCGGCGGCA 370  
QY 155 ACCGCGAGAACTACTATCTGGGGTGGCTGATCGAGAGGCGGCGTGGCGCAAGA 214  
DB 155 ACCGCGAGAACTACTATCTGGGGTGGCTGATCGAGAGGCGGCGTGGCGCAAGA 214  
QY 371 AGGTGAGCGAGAACTACTCCATCTGCGAGACCTGCGAGGCGGAGTGTGACAGGCGCA 430  
DB 371 AGGTGAGCGAGAACTACTCCATCTGCGAGACCTGCGAGGCGGAGTGTGACAGGCGCA 430  
QY 215 TCGACTTCGGGAGGCGCGGCGTGGCGTCTGGAGCGGCTGCTGAGGAGGAGGCGCA 274  
DB 215 TCGACTTCGGGAGGCGCGGCGTGGCGTCTGGAGCGGCTGCTGAGGAGGAGGCGCA 274  
QY 431 TCAGCCCGCCGACCTTGACCCCTGGGTGAAGGTGATCGAGGAGAGGCGCTTCAGGCGCG 490  
DB 431 TCAGCCCGCCGACCTTGACCCCTGGGTGAAGGTGATCGAGGAGAGGCGCTTCAGGCGCG 490  
QY 275 ACCAGCGGACCTTCGACTTCCTCTGCTGGAGCGCGGACAAAGGACAACTACCTCAACTAC 334  
DB 275 ACCAGCGGACCTTCGACTTCCTCTGCTGGAGCGCGGACAAAGGACAACTACCTCAACTAC 334  
QY 491 AGGTATCCCATCTTCACCCCTGAGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 520  
DB 491 AGGTATCCCATCTTCACCCCTGAGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 520  
QY 335 ACAGCGGCTCATGAAGTCTGTAAGGTGCGGCGGCTGCTGCGGTACGAC---AATAGCG 391  
DB 335 ACAGCGGCTCATGAAGTCTGTAAGGTGCGGCGGCTGCTGCGGTACGAC---AATAGCG 391  
QY 551 TTTTGAACACCTGTGGCGCGGCGGCGGCTGCGAGATGCTGAAGGAGGAGGCGGCGGCGGCA 610  
DB 551 TTTTGAACACCTGTGGCGCGGCGGCGGCTGCGAGATGCTGAAGGAGGAGGCGGCGGCGGCA 610  
QY 392 TCTGGAAGGCTCCGCTGCTGCTCCCGCGGAGCGCCCATGCGGCAAGTATCTGCTACT 451  
DB 392 TCTGGAAGGCTCCGCTGCTGCTCCCGCGGAGCGCCCATGCGGCAAGTATCTGCTACT 451  
QY 611 AGGAGGCGCGGAGTGGGAGCGGCTGTACCCGCTGACCCGCTGAGGCGGCG---CCGCTGGCGTCCG 667  
DB 611 AGGAGGCGCGGAGTGGGAGCGGCTGTACCCGCTGACCCGCTGAGGCGGCG---CCGCTGGCGTCCG 667  
QY 452 ACCGCGACTCTGCTCGAGTCAACAAGGCGCTCGCGCGGAGCGGCGTGGAGATCT 511  
DB 452 ACCGCGACTCTGCTCGAGTCAACAAGGCGCTCGCGCGGAGCGGCGTGGAGATCT 511  
QY 668 GCCAGATGGGACCGCGCGGCGGAGCATCGGGGCGGCGCA 712  
DB 668 GCCAGATGGGACCGCGCGGCGGAGCATCGGGGCGGCGCA 712  
QY 512 GCCAGCTCCCGTGGGACGGCATCACCTCTTCCCGCGCGGCGCA 556  
DB 512 GCCAGCTCCCGTGGGACGGCATCACCTCTTCCCGCGCGGCGCA 556

# RESULT 15

BJ261281.1

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshin@genes.nig.ac.jp.

Location/Qualifiers

1..694

/organism="Triticum aestivum"

/mol\_type="mRNA"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="whh2e03"

/tissue\_type="spike at heading date"

/dev\_stage="Peekes' scale 10.5"

/clone\_lib="Y. Ogihara unpublished cDNA library, wh\_h"

128 a 164 c 256 g 145 t 1 others

BASE COUNT

ORIGIN

Query Match 4.9%; Score 73.2; DB 12; Length 694;

Best Local Similarity 50.9%; Pred. No. 0.0085;

Matches 227; Conservative 0; Mismatches 213; Indels 6; Gaps 2;

QY 270 CGGCATCAGGTCCGCGACACCAAGGAGGCGCTGACAAAGATCGAGGAGGAGGAGCA 329

DB 270 CGGCATCAGGTCCGCGACACCAAGGAGGCGCTGACAAAGATCGAGGAGGAGGAGCA 329

QY 600 CGCCATCCCCCAGCAGCGCACCATCTTGGCCATGGCATCAACCGCAGAACTACGAGCT 541

DB 600 CGCCATCCCCCAGCAGCGCACCATCTTGGCCATGGCATCAACCGCAGAACTACGAGCT 541

QY 340 GTCCAGCAGAGAGACCCAGCAGGCGCAAGGAGGCGGCAAGGTGAGCCAGAACTACCC 369

DB 340 GTCCAGCAGAGAGACCCAGCAGGCGCAAGGAGGCGGCAAGGTGAGCCAGAACTACCC 369

QY 540 GGGGCTGCGGTGCATCGAGAGGCGGCGTGGCGCAAGATCGACTTCGCGAGGCGCC 481

DB 540 GGGGCTGCGGTGCATCGAGAGGCGGCGTGGCGCAAGATCGACTTCGCGAGGCGCC 481

QY 390 CATCTGCAAGAACCTGCGAGGCGGAGATGTCACAGGCGCATCAGCCCGCAGCCGTGA 449

DB 390 CATCTGCAAGAACCTGCGAGGCGGAGATGTCACAGGCGCATCAGCCCGCAGCCGTGA 449

QY 480 GCGCTGCGGCTGCTGGAGCGGCTGCTGGAGGAGGAGGCGCAACACCGGACCTTCGACTT 421

DB 480 GCGCTGCGGCTGCTGGAGCGGCTGCTGGAGGAGGAGGCGCAACACCGGACCTTCGACTT 421

QY 450 CGCCTGGGTGAAGGTGATCGAGGAGAGGCGCTTCAGGCGCGGAGGTGATCCCATGTTTAC 509

DB 450 CGCCTGGGTGAAGGTGATCGAGGAGAGGCGCTTCAGGCGCGGAGGTGATCCCATGTTTAC 509

QY 420 CGTCTTCGTGGAGCGCGCAAGGACAACTACCTCAACTACCGAGCGGCTCATGAAGCT 361

DB 420 CGTCTTCGTGGAGCGCGCAAGGACAACTACCTCAACTACCGAGCGGCTCATGAAGCT 361

QY 510 CGCCTTGAGGAGGCGCGCCACCCCGCAGGAGCCTGAACAGCATGTTGAACACCGTGGCGG 569

DB 510 CGCCTTGAGGAGGCGCGCCACCCCGCAGGAGCCTGAACAGCATGTTGAACACCGTGGCGG 569

QY 360 CUTCAGCTCGGCGGCTCCTCGGCTAGAC---AACACGCTTGGACAGGCTCCGCTCGT 304

DB 360 CUTCAGCTCGGCGGCTCCTCGGCTAGAC---AACACGCTTGGACAGGCTCCGCTCGT 304

QY 570 GCACGAGCGCGCATGCGAGATGCTGAAGSACACCATCAACGAGGAGGCGCGGAGTGGGA 629

DB 570 GCACGAGCGCGCATGCGAGATGCTGAAGSACACCATCAACGAGGAGGCGCGGAGTGGGA 629

QY 303 GCTCCCGCGGAGCGCCCATGCGCAAGTACATCGGCTACTACCGGCACTTCGCTCTGA 244

DB 303 GCTCCCGCGGAGCGCCCATGCGCAAGTACATCGGCTACTACCGGCACTTCGCTCTGA 244

QY 630 CGCGCTGACCCCGTGCAGGCGGCG---CCGCTGCGCGCGCGGCGGAGATGCGGAGCCCGG 686

DB 630 CGCGCTGACCCCGTGCAGGCGGCG---CCGCTGCGCGCGCGGCGGAGATGCGGAGCCCGG 686

QY 243 CACTCAACAGGCGCTCGCGCGCGGAGGAGGCGGCTGAGATCTGCCAGCTCCCGCTGGCGGA 184

DB 243 CACTCAACAGGCGCTCGCGCGCGGAGGAGGCGGCTGAGATCTGCCAGCTCCCGCTGGCGGA 184

QY 687 CGGAGGAGCATCGCGGCGGCGGCA 712

DB 687 CGGAGGAGCATCGCGGCGGCGGCA 712

QY 183 CGGCATCACCTCTGCGCGCGGCGCA 158

DB 183 CGGCATCACCTCTGCGCGCGGCGCA 158

Search completed: September 22, 2003, 22:50:20  
Job time : 3445.15 secs

	1	1509	100.0	1509	21	AAA15610	HIV synthetic Gag
2	1509	100.0	1509	24	AAI44549	HIV-1 p55gag poly(p)	
3	1509	100.0	1509	24	ABL39954	Synthetic Gag poly	
4	1491.4	98.8	1509	24	ABL39958	Synthetic Gag poly	
5	1489.8	98.7	1509	21	AAA15626	HIV codon-optimize	
6	1490.8	98.7	1509	24	AAI44553	HIV-1 p55gag poly(p)	
7	1297.6	86.0	1494	24	ABL35972	Synthetic Gag poly	
8	1294.8	85.8	1491	24	ABL40020	Synthetic Gag poly	

PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV  
PT Env polypeptide and the polypeptide useful for immunizing a mammal  
XX especially human against HIV  
PS  
PS Claim 2; Page 93; 113pp; English.  
XX  
CC Expression cassettes comprising a polynucleotide encoding antigenic  
CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are  
CC useful in DNA immunization, generation of packaging cell lines and  
CC production of Gag- and/or Env-containing proteins. Synthetic HIV and Gag  
CC expression cassettes exhibit increased potency for induction of  
CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1  
CC self-assemble into non-infectious virus-like particles which are used as  
CC a matrix for the proper presentation of an antigen entrapped or  
CC associated to the immune system of the host.  
XX  
SQ Sequence 1509 BP; 320 A; 556 G; 472 G; 161 T; 0 other;  
Query Match 100.0%; Score 1509; DB 21; Length 1509;  
Best Local Similarity 100.0%; Pred. No. 2.4e-179;  
Matches 1509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGCGCGCGCAGCATCTGCGGCGGAGAGCTGGACAGTGGGAGAGATCCGC 60  
DB 1 ATGGGCGCGCGCGCAGCATCTGCGGCGGAGAGCTGGACAGTGGGAGAGATCCGC 60  
QY 61 CTGCGCGCGCGCGCGCAAGACACTACATGCTGAAGCACTGGTGTGGCCAGCGCGAG 120  
DB 61 CTGCGCGCGCGCGCGCAAGACACTACATGCTGAAGCACTGGTGTGGCCAGCGCGAG 120  
QY 121 CTGGAGAGGCTTGGCGCTGAACCCCGGCTGTGTGAGACCGCCGAGGCTGTGACAGATC 180  
DB 121 CTGGAGAGGCTTGGCGCTGAACCCCGGCTGTGTGAGACCGCCGAGGCTGTGACAGATC 180  
QY 181 ATGAAGCACTGCAGCGCGCTGTGAGACCGGACCGAGGAGCTGCCAGCCTGTATCAAC 240  
DB 181 ATGAAGCACTGCAGCGCGCTGTGAGACCGGACCGAGGAGCTGCCAGCCTGTATCAAC 240  
QY 241 ACCGTGGCCACCTTACTTGGTGTGACCGCGGCTGCGAGTCCCGACACCAAGAGGCTC 300  
DB 241 ACCGTGGCCACCTTACTTGGTGTGACCGCGGCTGCGAGTCCCGACACCAAGAGGCTC 300  
QY 301 CTGGCAAGATCAGAGGAGCGAGAACAGTCCAGAGAGAGACCCAGCGGCGCAAGAG 360  
DB 301 CTGGCAAGATCAGAGGAGCGAGAACAGTCCAGAGAGAGACCCAGCGGCGCAAGAG 360  
QY 361 GCGCAGCGCAAGGTGAGCGAGCAACTACCCATGCTGCGAGAACTGCGAGGCGCAGATG 420  
DB 361 GCGCAGCGCAAGGTGAGCGAGCAACTACCCATGCTGCGAGAACTGCGAGGCGCAGATG 420  
QY 421 CACGAGGCGATCAGCGCGCGACCTGAACGCTGGGTGAAGTGAATCGAGGAGAGGCT 480  
DB 421 CACGAGGCGATCAGCGCGCGACCTGAACGCTGGGTGAAGTGAATCGAGGAGAGGCT 480  
QY 481 TTACGCGCGAGGTGATCCCGATGTCACCGCGCTGAGCGAGGCGGCGACCCCGCAGSAC 540  
DB 481 TTACGCGCGAGGTGATCCCGATGTCACCGCGCTGAGCGAGGCGGCGACCCCGCAGSAC 540  
QY 541 CTGACACCATGTTGAACACGCTGGCGGCGCACAGCGCGCCATGCGAGATGCTGAAGGAC 600  
DB 541 CTGACACCATGTTGAACACGCTGGCGGCGCACAGCGCGCCATGCGAGATGCTGAAGGAC 600  
QY 601 ACCATCAAGCAGGCGCGCGAGTGGGAGCGGCTGACCGCGCTGCGAGGCGGCGCGG 660  
DB 601 ACCATCAAGCAGGCGCGCGAGTGGGAGCGGCTGACCGCGCTGCGAGGCGGCGCGG 660  
QY 661 GCGCGCGCGCAGATGCGTGAACCGCGCGGCGAGAGACATGCGCGGCGCACAGACAGG 720  
DB 661 GCGCGCGCGCAGATGCGTGAACCGCGCGGCGAGAGACATGCGCGGCGCACAGACAGG 720  
QY 721 CAGGAGCAGATGCGCTGATGACAGCAACCGCGCGCTGCGCGTGGCGAGATCTCAAG 780  
DB 721 CAGGAGCAGATGCGCTGATGACAGCAACCGCGCGCTGCGCGTGGCGAGATCTCAAG 780

QY 761 CCGTGATATCTCTGGGCTGAACAAGATCTGCGGAAGTACAGCCCGGTGACATCTCTG 840  
DB 761 CCGTGATATCTCTGGGCTGAACAAGATCTGCGGAAGTACAGCCCGGTGACATCTCTG 840  
QY 841 GACATCGCGCAGGCGCGCCAGGAGCGCTTGGGAGCTACGTCAGCGCTCTTCAAGACC 900  
DB 841 GACATCGCGCAGGCGCGCCAGGAGCGCTTGGGAGCTACGTCAGCGCTCTTCAAGACC 900  
QY 901 CTGCGCGCGCAGGCGCGCCAGGAGCTGAAGAATGGATGACCGAGAGCCCTGCTGGT 960  
DB 901 CTGCGCGCGCAGGCGCGCCAGGAGCTGAAGAATGGATGACCGAGAGCCCTGCTGGT 960  
QY 961 CAGAACGCCAACCGGAGCTGCAAGCACTCTGGGCGCTCTCGGCGCGCGCGCCAGCTG 1020  
DB 961 CAGAACGCCAACCGGAGCTGCAAGCACTCTGGGCGCTCTCGGCGCGCGCGCCAGCTG 1020  
QY 1021 GAGGAGATGATGAGCGCTGCGAGGCGGTGGGCGCTCGGCGCAAGAGCGCGCTGCTG 1080  
DB 1021 GAGGAGATGATGAGCGCTGCGAGGCGGTGGGCGCTCGGCGCAAGAGCGCGCTGCTG 1080  
QY 1081 GCGAGGCGATGAGCGCGCGCAAGCGCTGAACATCATGATGAGAGAGCACTTCAAG 1140  
DB 1081 GCGAGGCGATGAGCGCGCGCAAGCGCTGAACATCATGATGAGAGAGCACTTCAAG 1140  
QY 1141 GCGCGCGCGCGCGCAAGCTCAAGTGTTCAGCTGGGCAAGGAGGCGCACATGCGCAAGAAC 1200  
DB 1141 GCGCGCGCGCGCGCAAGCTCAAGTGTTCAGCTGGGCAAGGAGGCGCACATGCGCAAGAAC 1200  
QY 1201 TGCGCGCGCGCGCGCAAGAGGCTGTGGAAATGCGGCAAGGAGGCGCACATGAG 1260  
DB 1201 TGCGCGCGCGCGCGCAAGAGGCTGTGGAAATGCGGCAAGGAGGCGCACATGAG 1260  
QY 1261 GACTGACCGAGCGCGCGCAACTTCTTGGGCAAGATCTGGCGCGCGCGCGCGCGCG 1320  
DB 1261 GACTGACCGAGCGCGCGCAACTTCTTGGGCAAGATCTGGCGCGCGCGCGCGCGCG 1320  
QY 1321 CCGCGCAACTTCTTGGCAAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
DB 1321 CCGCGCAACTTCTTGGCAAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
QY 1381 CCGCGCGAGAGCTTCCGCTTCAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440  
DB 1381 CCGCGCGAGAGCTTCCGCTTCAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440  
QY 1441 CCGGAGCGCTACCGCGAGCGCGCTGACCGCGCTGCGGCGCGCTGTTCGCGAGCGCGCG 1500  
DB 1441 CCGGAGCGCTACCGCGAGCGCGCTGACCGCGCTGCGGCGCGCTGTTCGCGAGCGCGCG 1500  
QY 1501 AGCCAGTAA 1509  
DB 1501 AGCCAGTAA 1509

RESULT 2  
AAL44549  
ID AAL44549 standard; DNA; 1509 BP.  
XX  
AC AAL44549;  
XX  
DT 06-NOV-2002 (first entry)  
XX  
DE HIV-1 p55gag polypeptide coding sequence 2.  
KW HIV; ds; vaccine; gene; immune response; microparticle;  
KW adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;  
KW polycaprolactone; polyorthoester; polycyanoacrylate; detergent;  
KW submicron emulsion; viral infection; bacterial infection;  
XX parasitic infection; HIV-1 p55gag polypeptide.  
CS Human immunodeficiency virus type 1.  
XX  
PN W0200226209-A2.

```

XX PD 04-APR-2002.
XX PF 28-SEP-2001; 2001KO-US30540.
XX PR 28-SEP-2000; 2000US-236105P.
XX PR 30-AUG-2001; 2001US-315905P.
XX PA (CHIR ) CHIRON CORP.
XX PI O'hagan D., Otten G., Donnelly JJ., Polo JM., Barnett S., Singh M.;
XX PI Ulmer J., Dubensky TW;
XX DR WPI: 2002-519084/55.
XX PT A microparticle to which a biologically active macromolecule is
XX PT adsorbed, for use as a vaccine composition to treat viral, bacterial or
XX PT parasitic infections, comprises a polymer microparticle, a detergent
XX PT and a submicron emulsion.
XX PS Claim 72; Fig 2; 100pp; English.
XX CC The invention relates to a method of raising an immune response in a host
XX CC animal. The method of the invention comprises administering a
XX CC microparticle that has an adsorbent surface to which a first biologically
XX CC active macromolecule (e.g. a nucleic acid) has been adsorbed. The
XX CC microparticle comprises a polymer microparticle of poly(alpha-hydroxy
XX CC acid), a polyhydroxy butyric acid, a polycaprolactone, a polycyanoester,
XX CC a polycyanacrylate, a detergent, and submicron emulsion. The method/
XX CC microparticle of the invention is useful for immunising a host animal
XX CC against viral, bacterial or parasitic infections. The present DNA
XX CC sequence encodes a HIV-1 p55gag polypeptide.
XX SO Sequence 1559 BP: 320 A: 556 C: 472 G: 161 T: 0 other;

```

```

Query Match      100.0%; Score 1509; DR 24; Length 1509;
Best Local Similarity 100.0%; Pred. No. 2.4e-179;
Matches 1509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 ATGGGCGCCCGCCAGCATCTCGCGGCGAGAGCTGGAAGTGGGAGAGATCGG 60
Db 1 ATGGGCGCCCGCCAGCATCTCGCGGCGAGAGCTGGAAGTGGGAGAGATCGG 60
Qy 61 CTGGGCGCCCGCGCGAGAGACTACATCTGATGACATCTGCTGGGCGGAGAG 120
Db 61 CTGGGCGCCCGCGCGAGAGACTACATCTGATGACATCTGCTGGGCGGAGAG 120
Qy 121 CTGGAGGGCTTGCCCTGACCCCGGCTCTGGAGACCGCGGAGGCTGCAAGGATC 180
Db 121 CTGGAGGGCTTGCCCTGACCCCGGCTCTGGAGACCGCGGAGGCTGCAAGGATC 180
Qy 181 ATGAAGCAGCTGAGCGCGCCCTGCGAGACCGCGGAGAGCTGCGAGCTGACAC 240
Db 181 ATGAAGCAGCTGAGCGCGCCCTGCGAGACCGCGGAGAGCTGCGAGCTGACAC 240
Qy 241 ACGGTGGCCACCTGACTGCTGTCAGCGCGGATCGAGTGGGACACCAAGAGGCC 300
Db 241 ACGGTGGCCACCTGACTGCTGTCAGCGCGGATCGAGTGGGACACCAAGAGGCC 300
Qy 301 CTGGACAAGATCAGGAGGAGAGAGAACTCCACGACGAGACCCAGCAGGCGAGAG 360
Db 301 CTGGACAAGATCAGGAGGAGAGAGAACTCCACGACGAGACCCAGCAGGCGAGAG 360
Qy 361 GCGACGCGCAAGGTGAGCGGAGAACTACCCCATCGTGCAGAACTCGAGGCGCAGTGTG 420
Db 361 GCGACGCGCAAGGTGAGCGGAGAACTACCCCATCGTGCAGAACTCGAGGCGCAGTGTG 420
Qy 421 CACGAGGCCATCAGCCCGCCACCTGAGCCCTGGTGAAGGTGATCAGAGAGAGGCC 480
Db 421 CACGAGGCCATCAGCCCGCCACCTGAGCCCTGGTGAAGGTGATCAGAGAGAGGCC 480
Qy 481 TTCAGCCCGAGGTGATCCCATGTTTCACCGGCTGAGGAGGGCGCCACCCCGAGGAC 540

```

RESULT 3  
ABL39954

```

Db 481 TTCAGCCCGAGGTGATCCCATGTTTACCGGCTTGAGCGAGGGGCGCCACCCCGAGGAC 540
Qy 541 CTGAACACAGTGTGAACACCGTGGCGGCCACCAAGTGGCCATGCGATGCTGAAGGAC 600
Db 541 CTGAACACAGTGTGAACACCGTGGCGGCCACCAAGTGGCCATGCGATGCTGAAGGAC 600
Qy 601 ACCATCAACGAGAGGCGCGCGAGTGGGACCGCTTCACCCCGTGCAGGCGGCGCCCGTG 660
Db 601 ACCATCAACGAGAGGCGCGCGAGTGGGACCGCTTCACCCCGTGCAGGCGGCGCCCGTG 660
Qy 661 GCGCCGCGCAGATGCGCGGACCGCCCGCGACGACATCGCGGCGCCACCAAGGACCGTG 720
Db 661 GCGCCGCGCAGATGCGCGGACCGCCCGCGACGACATCGCGGCGCCACCAAGGACCGTG 720
Qy 721 CAGGAGCAGATCGCTGGATGACCAACCGCCCGTGGCGGACATCTTACAAAG 780
Db 721 CAGGAGCAGATCGCTGGATGACCAACCGCCCGTGGCGGACATCTTACAAAG 780
Qy 781 CGGTGGATCATCTGGCGCTGACAGATCTGCGGATGACAGCCCGTGAAGATCTG 840
Db 781 CGGTGGATCATCTGGCGCTGACAGATCTGCGGATGACAGCCCGTGAAGATCTG 840
Qy 841 GACATCGCGCAGGCGCCCAAGGAGCGCTTCCGCGACTAGCTGAGCGCTTCTTCAAGACC 900
Db 841 GACATCGCGCAGGCGCCCAAGGAGCGCTTCCGCGACTAGCTGAGCGCTTCTTCAAGACC 900
Qy 901 CTGGGCGCGCAGGAGCGCCACCGAGAGCTGAGAGACTGATGACGAGACCGCTGTGTTG 960
Db 901 CTGGGCGCGCAGGAGCGCCACCGAGAGCTGAGAGACTGATGACGAGACCGCTGTGTTG 960
Qy 961 CAGAGCGCCAAACCGCGACTGCAAGACATCTTCCGCGCTCTCGGCGCGCGCGCCACCGTG 1020
Db 961 CAGAGCGCCAAACCGCGACTGCAAGACATCTTCCGCGCTCTCGGCGCGCGCGCCACCGTG 1020
Qy 1021 GAGAGATGATGACCGCTTCCAGGCGCTGGGCGCGCGCGCGCGCGCGCGCGCTGTG 1080
Db 1021 GAGAGATGATGACCGCTTCCAGGCGCTGGGCGCGCGCGCGCGCGCGCGCGCTGTG 1080
Qy 1081 GCGAGCGGATGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTG 1140
Db 1081 GCGAGCGGATGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTG 1140
Qy 1141 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
Db 1141 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
Qy 1201 TGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
Db 1201 TGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
Qy 1261 GACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
Db 1261 GACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
Qy 1321 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1321 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Qy 1381 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 1381 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Qy 1441 GCGAGCGCGCTACCGCGAGCGCGCTGACCGCGCTGCGGAGCGCTGTTCGGGAGCGCGCGCGTG 1500
Db 1441 GCGAGCGCGCTACCGCGAGCGCGCTGACCGCGCTGCGGAGCGCTGTTCGGGAGCGCGCGCGTG 1500
Qy 1501 AGCAGTAA 1509
Db 1501 AGCAGTAA 1509

```





```

QY 1381 CCGCCCGGAGAGCTTCGGCTTGGAGGAGACACACCCCGGCCCCCAAGCAGGAGCCCAAGGAC 1440
DB 1381 CCGCCCGGAGAGCTTCGGCTTGGAGGAGACACACCCCGGCCCCCAAGCAGGAGCCCAAGGAC 1440
QY 1441 GCGGAGCCCTACCGGAGCCCTTACCGGCGCTGACCGGCTTGGCAGAGCTGTGCGCAGCGCCCTG 1500
DB 1441 GCGGAGCCCTTACCGGAGCCCTTACCGGCGCTGACCGGCGCTGCGCAGAGCTGTGCGCAGCGCCCTG 1500
QY 1501 AGCCAGTAA 1509
DB 1501 AGCCAGTAA 1509

RESULT 4
ID ABL39958 standard; DNA: 1509 BP.
XX ABL39958;
XX AC
XX AT
XX CT
XX GT
XX T
XX Synthetic Gag polynucleotide sequence SEQ ID NO:21.
XX Human immunodeficiency virus type C; antigenic HIV type C protein;
XX immunogenic; immunisation: gag; pol; vif; vpr; tat; rev; vpu; env; nef;
XX immunostimulant; gene therapy: gene; ds.
XX Human immunodeficiency virus type C.
XX Synthetic.
XX WO200204493-A2.
XX 17-JAN-2002.
XX 05-JUL-2000; 2000US-0610313.
XX (CHIR) CHIRON CORP.
XX (UYST-) UNIV STELLENBOSCH.
XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg RJ.
XX WPI; 2002-154920/20.
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful
XX in applications including DNA immunization, or generation of packaging
XX cell lines, particularly in: gene therapy
XX Example 1: Fig 6: 23pp: English.
XX The present invention describes expression cassettes comprising a
XX polynucleotide sequence encoding a polypeptide comprising immunogenic
XX HIV type C polypeptides. The expression cassettes comprise any of the
XX HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
XX Nef (1). (1) have immunostimulant activity and can be used in gene
XX therapy. The HIV type C polynucleotides are useful in applications
XX including DNA immunisation, generation of packaging cell lines, and
XX production of HIV type C proteins. The polynucleotides are particularly
XX useful in gene therapy and DNA immunisation applications. ABL39942 to
XX ABL40054 and ABL06204 to ABL06215 represent sequences used in the
XX exemplification of the present invention.
XX Sequence 1509 BP; 321 A; 560 C; 470 G; 158 T; 0 other;
XX
XX Query Match 98.8%; Score 1491.4; DB 24; Length 1509;
XX Best Local Similarity 99.3%; Pred. No. 3.7e-177;
XX Matches 1498; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 ATGGGCGCCGCGCCAGCATCTTGGCGGAGAGCTGGACAGTGGGAGAGATTCGCG 60
DB 1 ATGGGCGCCGCGCGCCAGCATCTTGGCGGAGAGCTGGACAGTGGGAGAGATTCGCG 60

```

```

QY 51 CTGCGCCCGGCGGCAAGCACTACATGCTGAAGCACTGCTGTGGCCAGCCCGGAG 120
DB 51 CTGCGCCCGGCGGCAAGCACTACATGCTGAAGCACTGCTGTGGCCAGCCCGGAG 120
QY 121 CTGAGGGCTTGGCCCTGGAACCCGCGCTGCTGGAGACCGCGAGGCTGCAAGCAATC 180
DB 121 CTGAGGGCTTGGCCCTGGAACCCGCGCTGCTGGAGACCGCGAGGCTGCAAGCAATC 180
QY 181 ATGAAGCAGCTGCAGCGCCGCTTGCAGACCGGACCGAGAGCTGCGCAGCTGTACAAC 240
DB 181 ATGAAGCAGCTGCAGCGCCGCTTGCAGACCGGACCGAGAGCTGCGCAGCTGTACAAC 240
QY 241 ACCGTGCCACCCCTGTACTGCTGCGACGCGGCTATCAGAGTCCGCACCAAGAGGCC 300
DB 241 ACCGTGCCACCCCTGTACTGCTGCGACGCGGCTATCAGAGTCCGCACCAAGAGGCC 300
QY 301 CTGGACAAGATCGAGGAGGAGAGCAACAGTCCAGCAGAAGACCCAGGAGCCCAAGGAG 360
DB 301 CTGGACAAGATCGAGGAGGAGAGCAACAGAGCCAGCAGAGACCCAGGAGCCCAAGGAG 360
QY 361 GCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGAGGCGGCGAGATGGTG 420
DB 361 GCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGAGGCGGCGAGATGGTG 420
QY 421 CACGAGCCATCAGCCCGCCGACCCCTGCAAGCGCTGGGTGAAGTGTATCAGAGAGAGGCC 480
DB 421 CACGAGCCATCAGCCCGCCGACCCCTGCAAGCGCTGGGTGAAGTGTATCAGAGAGAGGCC 480
QY 481 TTCAGCCCGGAGTGTCCCATGTTTACCGCCCTGAGGAGGCGGCCACCCCGCCAGGAC 540
DB 481 TTCAGCCCGGAGTGTCCCATGTTTACCGCCCTGAGGAGGCGGCCACCCCGCCAGGAC 540
QY 541 CTGAACACGATGTTGACACCGTGGCGGCCACCGCTGAGGAGGCGGCCATCGAGATGCTGAAGGAC 600
DB 541 CTGAACACGATGTTGACACCGTGGCGGCCACCGCTGAGGAGGCGGCCATCGAGATGCTGAAGGAC 600
QY 601 ACCATCAACGAGAGGCGCGCGAGTGGGAGCGCTGCACCCGCTGAGGCGGCCCGCTG 660
DB 601 ACCATCAACGAGAGGCGCGCGAGTGGGAGCGCTGCACCCGCTGAGGCGGCCCGCTG 660
QY 661 GCGCCCGGCGAGTGGCGGACCCCGCGGAGGAGATCGCGGCGGCCACCGAGCCCTG 720
DB 661 GCGCCCGGCGAGTGGCGGACCCCGCGGAGGAGATCGCGGCGGCCACCGAGCCCTG 720
QY 721 CAGGACGAGATGCGCTGGATGATGACCAACCCCGCTGCGCGGCGGAGATCTACAAG 780
DB 721 CAGGACGAGATGCGCTGGATGATGACCAACCCCGCTGCGCGGCGGAGATCTACAAG 780
QY 781 CGGTGGATCATCTGGGCGCTGAACAGATCGTGGGATGTACAGCCCGCTGAGCATCTCTG 840
DB 781 CGGTGGATCATCTGGGCGCTGAACAGATCGTGGGATGTACAGCCCGCTGAGCATCTCTG 840
QY 841 GACATCGCGGAGGCGCCCAAGGAGCCCTTCGCGGATAGTGGAGCCGCTTCAAGAGCC 900
DB 841 GACATCGCGGAGGCGCCCAAGGAGCCCTTCGCGGATAGTGGAGCCGCTTCAAGAGCC 900
QY 901 CTGCGCGCGGAGGAGCCAGCCAGGAGCTGAAGAAGTGGATGACCGAGAGCCCTGCTGGTG 960
DB 901 CTGCGCGCGGAGGAGCCAGCCAGGAGCTGAAGAAGTGGATGACCGAGAGCCCTGCTGGTG 960
QY 961 CAGAACCCCAACCCGAGCTGCAAGACCCATCTCGCGGCTCTCGGCGCCGCGCCCGCTG 1020
DB 961 CAGAACCCCAACCCGAGCTGCAAGACCCATCTCGCGGCTCTCGGCGCCGCGCCCGCTG 1020
QY 1021 GAGGAGATGATGACCGGCTGCCAGGCGTGGCGGCGCGCGGCGCACAGCCCGCTGCTG 1080
DB 1021 GAGGAGATGATGACCGGCTGCCAGGCGTGGCGGCGCGCGGCGCACAGCCCGCTGCTG 1080
QY 1081 GCGGAGCGATGAGCCAGGCGCAACAGCGTGAACATCATGATGAGAGAGCAACTTCAAG 1140
DB 1081 GCGGAGCGATGAGCCAGGCGCAACAGCGTGAACATCATGATGAGAGAGCAACTTCAAG 1140
QY 1141 GGCCCCCGGCGGCAAGCTCAAGTCTTCACTGCGGCAAGGAGGCGCCACATCGCCAGAGAC 1200

```



```

|||||
Db 963 CAGAACCCCAACCCGACACACAGACATCTCTCGCGCCCTGGGCCCGCGGCACCTTC 1020
QY 1021 GAGGAGATGATACCGCTGCTCCAGGGCGTGGGGCGCCCGGCACAAAGCCGCGTGTG 1080
Db 1021 GAGGAGATGATACCGCTGCTCCAGGGCGTGGGGCGCCCGGCACAAAGCCGCGTGTG 1080
QY 1081 GCGGAGCGATAGCCAGGCGCACACCGTGAACATCATGATGAGAGAGCAACTTCAG 1140
Db 1081 GCGGAGCGATAGCCAGGCGCACACCGTGAACATCATGATGAGAGAGCAACTTCAG 1140
QY 1141 GGCCTCCGCGCCCAAGGTCAGTCTCACTGGGCAAGGAGGCGCACATGCCAAGAAC 1200
Db 1141 GGCCTCCGCGCCCAAGGTCAGTCTCACTGGGCAAGGAGGCGCACATGCCAAGAAC 1200
QY 1201 TGCGCGCCCGCCCGCAAGAGGCTGCTGGAATGCGGCAAGGAGGCGCACATGAG 1260
Db 1201 TGCGCGCCCGCCCGCAAGAGGCTGCTGGAATGCGGCAAGGAGGCGCACATGAG 1260
QY 1261 GACTGCACCGAGCGCCAGGCGCAACTTCTCTGGGCAAGATCTGSCCCAGCCACAGGCGCC 1320
Db 1261 GACTGCACCGAGCGCCAGGCGCAACTTCTCTGGGCAAGATCTGSCCCAGCCACAGGCGCC 1320
QY 1321 CCGCGCAACTTCTGCAAGAGCGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1321 CCGCGCAACTTCTGCAAGAGCGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
QY 1381 CCGCGCGAGAGCTTCCGCTTCGAGGAGACACCCCGCGCGCGCGCGCGCGCGCG 1440
Db 1381 CCGCGCGAGAGCTTCCGCTTCGAGGAGACACCCCGCGCGCGCGCGCGCGCGCG 1440
QY 1441 CCGGAGCGCTACCGCGAGCGCGCTGACGCGCGCTGCGGAGCGTGTCGCGAGCGCGCGCG 1500
Db 1441 CCGGAGCGCTACCGCGAGCGCGCTGACGCGCGCTGCGGAGCGTGTCGCGAGCGCGCGCG 1500
QY 1501 AGCCAGTAA 1509
Db 1501 AGCCAGTAA 1509

RESULT 6
AAL44553
ID AAL44553 standard: DNA: 1509 BP.
AC
XX
XX
XX 08-NOV-2002 (first entry)
XX
DE HIV-1 p55gag polypeptide coding sequence 4.
XX
KW HIV: ds; vaccine; gene; immune response; microparticle;
KW adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
KW polycaprolactone; polyorthoester; polycyanoacrylate; detergent;
KW submicron emulsion; viral infection; bacterial infection;
KW parasitic infection; HIV-1 p55gag polypeptide.
XX
OS Human immunodeficiency virus type 1.
XX
PH Key
FT variation Location/Qualifiers
FT replace (282, C)
FT /*tag= a
FT variation replace (331, T)
FT /*tag= b
FT variation replace (332, C)
FT /*tag= c
FT variation replace (549, G)
FT /*tag= d
FT variation replace (553, T)
FT /*tag= e
FT variation replace (783, G)
FT /*tag= f
FT variation replace (816, G)
FT /*tag= g

```

```

variation replace (999, T)
FT /*tag= h
FT variation replace (1002, C)
FT /*tag= i
FT variation replace (1089, G)
FT /*tag= j
FT variation replace (1149, G)
FT /*tag= k
FT variation replace (1158, C)
FT /*tag= l
XX W0200226209-A2.
XX
XX 04-APR-2002.
XX
XX 28-SEP-2001; 2001WO-US30540.
XX
XX 28-SEP-2000; 2000US-236105P.
XX 30-AUG-2001; 2001US-315905P.
XX (CHIR ) CHIRON CORP.
XX
XX O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;
XX Ulmer J, Dubensky TW;
XX WPI; 2002-519084/55.
XX
XX A microparticle to which a biologically active macromolecule is
XX adsorbed, for use as a vaccine composition to treat viral, bacterial or
XX parasitic infections, comprises a polymer microparticle, a detergent
XX and a submicron emulsion.
XX
XX Claim 72; Fig 6; 100pp; English.
XX
XX The invention relates to a method of raising an immune response in a host
XX animal. The method of the invention comprises administering a
XX microparticle that has an adsorbent surface to which a first biologically
XX active macromolecule (e.g. a nucleic acid) has been adsorbed. The
XX microparticle comprises a polymer microparticle of poly(alpha-hydroxy
XX acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
XX a polycyanoacrylate, a detergent, and submicron emulsion. The method/
XX microparticle of the invention is useful for immunising a host animal
XX against viral, bacterial or parasitic infections. The present DNA
XX sequence encodes a HIV-1 p55gag polypeptide.
XX
XX Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 other:
XX
Query Match 98.7%; Score 1489.8; DB 24; Length 1509;
Best Local Similarity 99.2%; Pred. No. 5.8e-177; Mismatches 12; Indels 0; Gaps 0;
Matches 1497; Conservative 0;
QY 1 ATGGGCGCGCGCGCAGCATCTCTCGCGGGGAGAAAGCTGGCAAGTGGGAGAGATCCGC 60
Db 1 ATGGGCGCGCGCGCAGCATCTCTCGCGGGGAGAAAGCTGGCAAGTGGGAGAGATCCGC 60
QY 61 CTGGCGCGCGCGCGGCAAGAACACTACATGCTGAAGCACTGTGTGGGCGACCGCGGAG 120
Db 61 CTGGCGCGCGCGCGGCAAGAACACTACATGCTGAAGCACTGTGTGTGGGCGACCGCGGAG 120
QY 121 CTGGAGGCGCTTCGCCCTGCAACCCCGCGGCTGCTGAGAGACCGCGGAGGCTGCAAGCAGATC 180
Db 121 CTGGAGGCGCTTCGCCCTGCAACCCCGCGGCTGCTGAGAGACCGCGGAGGCTGCAAGCAGATC 180
QY 181 ATGAAGCAGCTGCGAGCGCGCGCTGCGAGACCGGCGCTGCGAGAGCTGGCGAGCTGTCAAC 240
Db 181 ATGAAGCAGCTGCGAGCGCGCGCTGCGAGACCGGCGCTGCGAGAGCTGGCGAGCTGTCAAC 240
QY 241 ACCGTGGCCACCCCTGTACTGGTGCACCGCGGCGATCCAGTCCGCGACACCAAGAGGCC 300
Db 241 ACCGTGGCCACCCCTGTACTGGTGCACCGCGGCGATCCAGTCCGCGACACCAAGAGGCC 300
QY 301 CTGGACAAAGATCGGAGGAGGACAGCAAGTCCCGAGAGAGACCCAGCAGGCGCAAGGAG 360
Db 301 CTGGACAAAGATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360

```

```
Db 301 CTGACAAAGATCGAGGAGGACGACAAACAGAGGCGCAGAGAGACCCAGCAGGCGCAAGGAG 360
Qy 361 GCCAGCGCAAGGTGAGCGAGAACTACCCATCGTCAGAAACCTGCGAGGGCGCAGATGGTG 420
Db 361 GCCAGCGCAAGGTGAGCGAGAACTACCCATCGTCAGAAACCTGCGAGGGCGCAGATGGTG 420
Qy 421 CACGAGGCCATCAGCCCGGCGACGCTGACGCCCTGGGTGGAAGGTGATCGAGGAGAAAGCC 480
Db 421 CACGAGGCCATCAGCCCGGCGACGCTGACGCCCTGGGTGGAAGGTGATCGAGGAGAAAGCC 480
Qy 481 TTCAGCGCGAGGTGATCGCATGTTCAGCGCCCTGAGCGAGGCGGCGACCCCGCCGAG 540
Db 481 TTCAGCGCGAGGTGATCGCATGTTCAGCGCCCTGAGCGAGGCGGCGACCCCGCCGAG 540
Qy 541 CTGAACACGATGTGACACCGCTGGCGGCGCACAGCGCGCATGACAGATCTGAAAGGAC 600
Db 541 CTGAACACGATGTGACACCGCTGGCGGCGCACAGCGCGCATGACAGATCTGAAAGGAC 600
Qy 601 ACCATCAAGAGGAGCGCGCGAGTGGGACCGCTGACAGCCGCTGCAGGCGCGGCGGCTG 660
Db 601 ACCATCAAGAGGAGCGCGCGAGTGGGACCGCTGACAGCCGCTGCAGGCGCGGCGGCTG 660
Qy 661 CCCCCCGCCAGATGCGCGACCGCGCGGCGCAGCATCGCGCGCGCGCAGCAGCATG 720
Db 661 CCCCCCGCCAGATGCGCGACCGCGCGGCGCAGCATCGCGCGCGCGCAGCAGCATG 720
Qy 721 CAGGAGCAGATCGCGTGGATCAGCAGCAACCCCGCGTSCCGGTGGCGCAGCATACAG 780
Db 721 CAGGAGCAGATCGCGTGGATCAGCAGCAACCCCGCGTGGCGGTGGCGCAGCATACAG 780
Qy 781 CGGTGGATCATCTGGCGCTGAACAGATCGTGGGATGTACAGCCCGTGGAGCATCTG 840
Db 781 CGGTGGATCATCTGGCGCTGAACAGATCGTGGGATGTACAGCCCGTGGAGCATCTG 840
Qy 841 GACATCCCGCAGGCGCCCAAGAGCGCTTCGCGGACTACGTGGAGCGGCTTCTCAAGACC 900
Db 841 GACATCCCGCAGGCGCCCAAGAGCGCTTCGCGGACTACGTGGAGCGGCTTCTCAAGACC 900
Qy 901 CTGCGCGCGCAGGAGCCACCCAGACGCTGAAGAACTGGATGACGAGACCGCTGCTG 960
Db 901 CTGCGCGCGCAGGAGCCACCCAGACGCTGAAGAACTGGATGACGAGACCGCTGCTG 960
Qy 961 CAGACGCGCAACCCGAGCTGCAAGACCTCTGCGCGCTCTGCGCCCGCGGCGCACCTG 1020
Db 961 CAGACGCGCAACCCGAGCTGCAAGACCTCTGCGCGCTCTGCGCCCGCGGCGCACCTG 1020
Qy 1021 CAGGAGATGATGACCGCTGCGAGGCGCTGGCGGCGCGGCGCACAGCGCGGCTGCTG 1080
Db 1021 CAGGAGATGATGACCGCTGCGAGGCGCTGGCGGCGCGGCGCACAGCGCGGCTGCTG 1080
Qy 1081 GCGGAGCGATGAGCGAGGCGCAACAGCGTGAACATCATGTGAGAGAGCACTTCAAG 1140
Db 1081 GCGGAGCGATGAGCGAGGCGCAACAGCGTGAACATCATGTGAGAGAGCACTTCAAG 1140
Qy 1141 GSCCGCGCGCGCAAGTCAAGTGTCTCAACTGGGCAAGGAGGCGCAGATCGCAAGGAC 1200
Db 1141 GSCCGCGCGCGCAAGTCAAGTGTCTCAACTGGGCAAGGAGGCGCAGATCGCAAGGAC 1200
Qy 1201 TGCCCGCGCGCGCGCAAGTGTCTGGAAGTGGCGCAAGGAGGCGCAGATCGCAAGGAC 1260
Db 1201 TGCCCGCGCGCGCGCAAGTGTCTGGAAGTGGCGCAAGGAGGCGCAGATCGCAAGGAC 1260
Qy 1261 GACTGCACTGAGCGGCGCAGCGCAACTTCTGGGCAAGATCTGGGCGCAGCGCAAGGCGG 1320
Db 1261 GACTGCACTGAGCGGCGCAGCGCAACTTCTGGGCAAGATCTGGGCGCAGCGCAAGGCGG 1320
Qy 1321 CCGCGCAACTTCTGCAAGAGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1321 CCGCGCAACTTCTGCAAGAGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Qy 1381 CCGCGCGAGAGCTTCCGCTTCGAGGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCG 1440
Db 1381 CCGCGCGAGAGCTTCCGCTTCGAGGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCG 1440
```

```
Qy :441 CGGAGCCCTACCGGAGCGCCCTGACGCCCTGCGAGCCTGCGAGCGCGCCCTG 1500
Db :441 CGGAGCCCTACCGGAGCGCCCTGACGCCCTGCGAGCCTGCGAGCGCGCCCTG 1500
Qy :501 AGCCAGTAA 1509
Db :501 AGCCAGTAA 1509

RESULT 7
ABL39972
ID ABL39972 standard: DNA: 1494 BP.
XX AC ABL39972:
XX CT 15-MAY-2002 (first entry)
XX XX Synthetic Gag polynucleotide sequence SEQ ID NO:51.
XX KW Human immunodeficiency virus type C; antigenic HIV type C protein;
XX KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
XX KW immunostimulant; gene therapy; gene; ds.
XX CS Human immunodeficiency virus type C.
XX GS Synthetic.
XX WO200204493-A2.
XX EN 17-JAN-2002.
XX PD 05-JUL-2001; 2001WO-US21241.
XX PF 05-JUL-2000; 2000US-0610313.
XX PR (CHIR ) CHIRON CORP.
XX PA (UYST-) UNIV STELLENBOSCH.
XX FA Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX FI WPI: 2002-154920/20.
XX DR New polynucleotides encoding antigenic HIV Type C polypeptides, useful
XX PT in applications including DNA immunization or generation of packaging
XX PT cell lines, particularly in gene therapy -
XX PS Claim 7; Fig 22; 23pp; English.
XX CC The present invention describes expression cassettes comprising a
XX CC polynucleotide sequence encoding a polypeptide comprising immunogenic
XX CC HIV type C polypeptides. The expression cassettes comprise any of the
XX CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
XX CC Nef (i). (i) have immunostimulant activity and can be used in gene
XX CC therapy. The HIV type C polynucleotides are useful in applications
XX CC including DNA immunisation, generation of packaging cell lines, and
XX CC production of HIV type C proteins. The polynucleotides are particularly
XX CC useful in gene therapy and DNA immunisation applications. ABL39942 to
XX CC ABL40054 and ABB06204 to ABB06215 represent sequences used in the
XX CC exemplification of the present invention.
XX SQ Sequence 1494 BP; 325 A; 557 C; 455 G; 157 T; 0 other;

Query Match 86.0%; Score 1297.6; DB 24; Length 1494;
Best Local Similarity 92.7%; Pred. No. 4,le-153;
Matches 1402; Conservative 0; Mismatches 89; Indels 21; Gaps 3;
```

```
Qy 1 ATGGGCGCGCGCGCCATCTTCCGCGGCGGAGAGCTGGACAACTGGGAGAGATCCGC 60
Db 1 ATGGGCGCGCGCGCCATCTTCCGCGGCGGAGAGCTGGACAACTGGGAGAGATCCGC 60
Qy 61 CTGGCGCGCGCGCGGCAAGAGCACTACATCTGAAGCACCTGTGTGGCGCAGCCGCGAG 120
Db 61 CTGGCGCGCGCGCGGCAAGAGCACTACATCTGAAGCACCTGTGTGTGGCGCGAG 120
```

QY 121 CTGGAGCGCTTCCGCTGAACTCCGGCTGCTGGAGACCGCGAGGCGTGCAGACGATC 180  
DB 122 CTGGAGCGCTTCCGCTGAACTCCGGCTGCTGGAGACCGCGAGGCGTGCAGACGATC 180  
QY 181 ATGAAGCAGCTGAGCGCGCTGTCAGACCGGACCGAGGAGCTGCGAGAGCTGACAA 240  
DB 182 ATGAAGCAGCTGAGCGCGCTGTCAGACCGGACCGAGGAGCTGCGAGAGCTGACAA 240  
QY 241 ACCGTGGCCAGCTGACTGCTGTCAGACCGGACCGAGGAGCTGCGAGAGCTGACAA 300  
DB 242 ACCGTGGCCAGCTGACTGCTGTCAGACCGGACCGAGGAGCTGCGAGAGCTGACAA 300  
QY 301 CTGGACAAGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
DB 302 CTGGACAAGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
QY 361 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
DB 362 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
QY 421 CACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
DB 422 CACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
QY 481 TTCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
DB 482 TTCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
QY 541 CTGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
DB 542 CTGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
QY 601 ACCATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
DB 602 ACCATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
QY 661 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
DB 662 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
QY 721 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
DB 722 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
QY 781 CGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
DB 782 CGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
QY 841 GACATCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
DB 842 GACATCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
QY 901 CTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
DB 902 CTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
QY 961 CAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
DB 962 CAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
QY 1021 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
DB 1022 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
QY 1081 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
DB 1082 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
QY 1141 GCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
DB 1142 GCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
QY 1138 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1197  
DB 1139 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1197

QY 1201 TCGCGCGCGCGCGCGAGAGGCGTCTGCAAGTGGCGCAAGGAGGCGCCAGATGATGAG 1260  
DB 1202 TCGCGCGCGCGCGCGAGAGGCGTCTGCAAGTGGCGCAAGGAGGCGCCAGATGATGAG 1260  
QY 1257 TCGCGCGCGCGCGCGAGAGGCGTCTGCAAGTGGCGCAAGGAGGCGCCAGATGATGAG 1257  
DB 1258 TCGCGCGCGCGCGCGAGAGGCGTCTGCAAGTGGCGCAAGGAGGCGCCAGATGATGAG 1258  
QY 1261 GACTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
DB 1262 GACTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
QY 1256 GACTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1317  
DB 1257 GACTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1317  
QY 1321 CCGCGCAACTTCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1377  
DB 1322 CCGCGCAACTTCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1377  
QY 1318 CCGCGCAACTTCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1377  
DB 1319 CCGCGCAACTTCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1377  
QY 1378 CCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1437  
DB 1379 CCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1437  
QY 1378 CCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1437  
DB 1379 CCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1437  
QY 1438 GACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1497  
DB 1439 GACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1497  
QY 1438 -----CGCGAGCGCGCTGACGAGCGTGAAGAGCGTGTTCGGCGAGCGACCC 1482  
DB 1439 CTGAGCGCAGTAA 1509  
DB 1483 CTGAGCGCAGTAA 1494  
RESULT 6  
ABL40620  
ID ABL40620 standard; DNA; 1491 bp.  
XX ABL40620;  
AC ABL40620;  
DT 15-MAY-2002 (first entry)  
XX Synthetic Gag polynucleotide sequence SEQ ID NO:99.  
XX Human immunodeficiency virus type C; antigenic HIV type C protein;  
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; env; nef;  
XX immunostimulant; gene therapy; gene; ds.  
OS Human immunodeficiency virus type C.  
XX Synthetic.  
XX WC200204493-A2.  
XX 17-JAN-2002.  
XX 05-JUL-2001; 2001WO-US21241.  
XX 05-JUL-2000; 2000US-0610313.  
XX (CHIR) CHIRON CORP.  
XX (UYST-) UNIV STELLENBOSCH.  
PI Zur Megede J, Barnett SW, Engelbrecht S, Van Kensburg EJ,  
XX WPI; 2002-154920/20.  
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful  
PT in applications including DNA immunization or generation of packaging  
PT cell lines, particularly in gene therapy.  
XX Claim 8; Fig 70; 233pp; English.  
XX The present invention describes expression cassettes comprising a  
CC polynucleotide sequence encoding a polypeptide comprising immunogenic  
CC HIV type C polypeptides. The expression cassettes comprise any of the  
CC HIV type C sequences encoding gag, pol, vif, vpr, tat, rev, env or  
CC Nef (1). (1) have immunostimulant activity and can be used in gene  
CC therapy. The HIV type C polynucleotides are useful in applications  
CC including DNA immunisation, generation of packaging cell lines, and  
CC production of HIV type C proteins. The polynucleotides are particularly  
CC useful in gene therapy and DNA immunisation applications. ABL39942 to  
CC ABL40054 and ABL06204 to ABL06215 represent sequences used in the

exemplification of the present invention.

Sequence 1491 BP; 319 A; 557 C; 457 G; 158 T; 0 other;

```

every Match      85.8%; Score 1294.8; DB 24; Length 1491;
1st Local Similarity 92.7%; Pred. No. 9,1e-153;
Matches 1401; Conservative 0; Mismatches 87; Indels 24;

```

1 ATGGCGCGCGCGCAGAGCATCTCTGCGCGCGCGAGAAAGTGTGGACAAGTGGGAGAAAGATCGCG 60  
2 ATGGCGCGCGCGCGCAGAGCATCTCTGCGCGCGCGAGAGTGTGGACAAGTGGGAGAAAGATCGCG 60  
61 CTGGCGCGCGCGCGCAAGACACTACATCTCTGAAGACACCTGSGTGTGGCGCAACCGCGGAG 120  
61 CTGGCGCGCGCGCGCGCAAGACACTACATCTCTGAAGACACCTGSGTGTGGCGCAACCGCGGAG 120  
121 CTGGAGGCGCTTGGCGCTGAACCGCGCGCGCTCTGGAGACCGCGACCGAGAGCTGGCGAGCTGTATCAAC 240  
121 CTGGAGGCGCTTGGCGCTGAACCGCGCGCGCTCTGGAGACCGCGACCGAGAGCTGGCGAGCTGTATCAAC 240  
181 ATGAAGCAGCTGCAGCGCGCGCGCTGCGAGACCGCGACCGCGAGAGCTGGCGAGCTGTGTATCAAC 240  
181 ATCAAGCAGCTGCAGCGCGCGCGCTGCGAGACCGCGACCGCGAGAGCTGGCGAGCTGTGTATCAAC 240  
241 ACCGTGGCGCACCGCTGTATCTGGTGTGACCGCGCGCATCGAGGTCCCGCGACACCAAGGAGGCC 300  
241 ACCGTGGCGCACCGCTGTATCTGGTGTGACCAAGGCGATCGACGTGCCGACACCAAGGAGGCC 300  
301 CTGGACAAGATCGAGGAGGAGCAACAAGTCCGACGAGAAGCCCGACCGAGCCCAAGGAG 360  
301 CTGGACAAGATCGAGGAGGAGCAACAAGTCCGACGAGAAGATCCGACGAGAAGCCCGACCGAGGCC 360  
361 GCGCAGCGCAAGGTGAGCGCAAGACTACCCCATCTGTGAGACCTGCGAGGCGCGCATGCTG 420  
361 GCGCAGCAAGAGGTGAGCGCAAGACTACCCCATCTGTGAGAACCTGCGAGGCGCGCATGCTG 420  
421 CACGAGGCGCATCAGCGCGCGCGCACCTTGACCGCTGGGTGAAGGTGATCGAGGAGAGGCC 480  
421 CACGAGGCGCATCAGCGCGCGCGCACCTTGACCGCTGGGTGAAGGTGATCGAGGAGAGGCC 480  
481 TTCAGCGCGGAGGTATCCCATGTTTACCGCTCTGAGCGAGGCGCGCACCCCGCAGGAC 540  
481 TTCAGCGCGGAGGTATCCCATGTTTACCGCTCTGAGCGAGGCGCGCACCCCGCAGGAC 540  
541 CTGAACACGATGTTGAACACCGTGGCGCGCACCGAGCGCGCATGCGAGATGCTGAAGGAC 600  
541 CTGAACACGATGTTGAACACCGTGGCGCGCACCGAGCGCGCATGCGAGATGCTGAAGGAC 600  
601 ACCATCAAGGAGGCGCGGAGTGGGACCGCGCTGACCCCGCTGCGAGCGCGCGCGCGCTG 660  
601 ACCATCAAGGAGGCGCGGAGTGGGACCGCGCTGACCCCGCTGCGAGCGCGCGCGCGCTG 660  
661 GCGCGCGCGCAGATCGCGGACCGCGCGCGGAGCGGACATCGCGCGCGCGCGCACCGACCGCTG 720  
661 GCGCGCGCGCAGATCGCGGAGCGCGCGGAGCGGACATCGCGCGCGCGCGCACCGACCGCTG 720  
721 CAGGAGCAGATCGCTGGATGACCAAGCAACCGCGCGCTGCGCGTGGCGGACATCTACAA 780  
721 CAGGAGCAGATCGCTGGATGACCAAGCAACCGCGCGCTGCGCGTGGCGGACATCTACAA 780  
781 CGGTGGATCATGCTGGGCTGAACAAGATCGTGGGATGTACAGCGCGCGTGAAGTATCTG 840  
781 CGGTGGATCATGCTGGGCTGAACAAGATCGTGGGATGTACAGCGCGCGTGAAGTATCTG 840  
841 GACATCGCGCGCGCGCGCGCGCTTCCGCGCTACGTGGAGCGCTCTTCAAGAGCC 900  
841 GACATCAAGCAGGCGCGCGCGCGCGCTTCCGCGCTACGTGGAGCGCTCTTCAAGAGCC 900  
901 CTGCGCGCGCAGCAGGCGCGCGCGCGCGCTGATGAGCACTGGATGAGCGCGAGACCGCTGCTG 960  
901 CTGCGCGCGCAGCAGCAGCAGCGCGCGCGCGCGCTGATGAGCACTGGATGAGCGCGAGACCGCTGCTG 960  
961 CAGAAACCGCAACCGCGCTGCAAGACCATCTGCGCGCGCTCTCGCGCGCGCGCGCGCGCTG 1020



polycaprolactone; polyorthoester; polycyanoacrylate; detergent;  
submicron emulsion; viral infection; bacterial infection;  
parasitic infection; HIV-1 p53gag polypeptide.

submicron emulsion; viral infection; bacterial infection;

parasitic infection: HIV-1 p55 gag polypeptide.

0  
4  
8  
12  
16  
20  
24  
28  
32  
36  
40  
44  
48  
52  
56  
60  
64  
68  
72  
76  
80  
84  
88  
92  
96  
100

အသံအသွယ်

## DIFFICULT QUESTIONS

WQ200226209-A2

NOZ00240203 A2.

04-APR-2002

04-APR-2002.

0F50C351-0M160C : 160C-C35-8C

EC-DMT007 ; T007-435-07

79-SEP-2000: 0000UTC-

(CHIR ) CHIRON CORP.  
28-SEP-2000; 2000US-236193P.  
30-AUG-2001; 2000US-315905P.

O'hagan D, Otter G, Donnelly J

Ulmer J, Dubensky JW;  
WPI: 2002-519084/55.

A microparticle to which a biologically active macromolecule is adsorbed, for use as a vaccine composition to treat viral, bacterial or parasitic infections, comprises a polymer microparticle, a detergent and a submicron emulsion.

Claim 72: Fig 1: 100pp: English.

The invention relates to a method of raising an immune response in a host animal. The method of the invention comprises administering a microparticle that has an adsorbent surface to which a first biologically active macromolecule (e.g. a nucleic acid) has been adsorbed. The microparticle comprises a polymer microparticle of poly(alpha-hydroxy acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester, a polycyanacrylate, a detergent, and submicron emulsion. The method/ microparticle of the invention is useful for immunising a host animal against viral, bacterial or parasitic infections. The present DNA sequence encodes a HIV-1 p56gag polypeptide.

1. *Chlorophyll a* (Chl *a*)  
 2. *Chlorophyll b* (Chl *b*)  
 3. *Chlorophyll c* (Chl *c*)  
 4. *Chlorophyll d* (Chl *d*)  
 5. *Chlorophyll e* (Chl *e*)  
 6. *Chlorophyll f* (Chl *f*)  
 7. *Chlorophyll g* (Chl *g*)  
 8. *Chlorophyll h* (Chl *h*)  
 9. *Chlorophyll i* (Chl *i*)  
 10. *Chlorophyll j* (Chl *j*)  
 11. *Chlorophyll k* (Chl *k*)  
 12. *Chlorophyll l* (Chl *l*)  
 13. *Chlorophyll m* (Chl *m*)  
 14. *Chlorophyll n* (Chl *n*)  
 15. *Chlorophyll o* (Chl *o*)  
 16. *Chlorophyll p* (Chl *p*)  
 17. *Chlorophyll q* (Chl *q*)  
 18. *Chlorophyll r* (Chl *r*)  
 19. *Chlorophyll s* (Chl *s*)  
 20. *Chlorophyll t* (Chl *t*)  
 21. *Chlorophyll u* (Chl *u*)  
 22. *Chlorophyll v* (Chl *v*)  
 23. *Chlorophyll w* (Chl *w*)  
 24. *Chlorophyll x* (Chl *x*)  
 25. *Chlorophyll y* (Chl *y*)  
 26. *Chlorophyll z* (Chl *z*)  
 27. *Chlorophyll aa* (Chl *aa*)  
 28. *Chlorophyll ab* (Chl *ab*)  
 29. *Chlorophyll ac* (Chl *ac*)  
 30. *Chlorophyll ad* (Chl *ad*)  
 31. *Chlorophyll ae* (Chl *ae*)  
 32. *Chlorophyll af* (Chl *af*)  
 33. *Chlorophyll ag* (Chl *ag*)  
 34. *Chlorophyll ah* (Chl *ah*)  
 35. *Chlorophyll ai* (Chl *ai*)  
 36. *Chlorophyll aj* (Chl *aj*)  
 37. *Chlorophyll ak* (Chl *ak*)  
 38. *Chlorophyll al* (Chl *al*)  
 39. *Chlorophyll am* (Chl *am*)  
 40. *Chlorophyll an* (Chl *an*)  
 41. *Chlorophyll ao* (Chl *ao*)  
 42. *Chlorophyll ap* (Chl *ap*)  
 43. *Chlorophyll aq* (Chl *aq*)  
 44. *Chlorophyll ar* (Chl *ar*)  
 45. *Chlorophyll as* (Chl *as*)  
 46. *Chlorophyll at* (Chl *at*)  
 47. *Chlorophyll au* (Chl *au*)  
 48. *Chlorophyll av* (Chl *av*)  
 49. *Chlorophyll aw* (Chl *aw*)  
 50. *Chlorophyll ax* (Chl *ax*)  
 51. *Chlorophyll ay* (Chl *ay*)  
 52. *Chlorophyll az* (Chl *az*)  
 53. *Chlorophyll aza* (Chl *aza*)  
 54. *Chlorophyll abz* (Chl *abz*)  
 55. *Chlorophyll acz* (Chl *acz*)  
 56. *Chlorophyll adz* (Chl *adz*)  
 57. *Chlorophyll aez* (Chl *aez*)  
 58. *Chlorophyll afz* (Chl *afz*)  
 59. *Chlorophyll agz* (Chl *agz*)  
 60. *Chlorophyll ahz* (Chl *ahz*)  
 61. *Chlorophyll aiz* (Chl *aiz*)  
 62. *Chlorophyll ajz* (Chl *ajz*)  
 63. *Chlorophyll akz* (Chl *akz*)  
 64. *Chlorophyll alz* (Chl *alz*)  
 65. *Chlorophyll amz* (Chl *amz*)  
 66. *Chlorophyll anz* (Chl *anz*)  
 67. *Chlorophyll aoz* (Chl *aoz*)  
 68. *Chlorophyll apz* (Chl *apz*)  
 69. *Chlorophyll aqz* (Chl *aqz*)  
 70. *Chlorophyll arz* (Chl *arz*)  
 71. *Chlorophyll asz* (Chl *asz*)  
 72. *Chlorophyll atz* (Chl *atz*)  
 73. *Chlorophyll auz* (Chl *auz*)  
 74. *Chlorophyll avz* (Chl *avz*)  
 75. *Chlorophyll awz* (Chl *awz*)  
 76. *Chlorophyll axz* (Chl *axz*)  
 77. *Chlorophyll ayz* (Chl *ayz*)  
 78. *Chlorophyll ayz* (Chl *ayz*)  
 79. *Chlorophyll azz* (Chl *azz*)  
 80. *Chlorophyll azaa* (Chl *aza*)  
 81. *Chlorophyll abz* (Chl *abz*)  
 82. *Chlorophyll acz* (Chl *acz*)  
 83. *Chlorophyll adz* (Chl *adz*)  
 84. *Chlorophyll aez* (Chl *aez*)  
 85. *Chlorophyll afz* (Chl *afz*)  
 86. *Chlorophyll agz* (Chl *agz*)  
 87. *Chlorophyll ahz* (Chl *ahz*)  
 88. *Chlorophyll aiz* (Chl *aiz*)  
 89. *Chlorophyll ajz* (Chl *ajz*)  
 90. *Chlorophyll akz* (Chl *akz*)  
 91. *Chlorophyll alz* (Chl *alz*)  
 92. *Chlorophyll amz* (Chl *amz*)  
 93. *Chlorophyll anz* (Chl *anz*)  
 94. *Chlorophyll aoz* (Chl *aoz*)  
 95. *Chlorophyll apz* (Chl *apz*)  
 96. *Chlorophyll aqz* (Chl *aqz*)  
 97. *Chlorophyll arz* (Chl *arz*)  
 98. *Chlorophyll asz* (Chl *asz*)  
 99. *Chlorophyll atz* (Chl *atz*)  
 100. *Chlorophyll auz* (Chl *auz*)  
 101. *Chlorophyll avz* (Chl *avz*)  
 102. *Chlorophyll awz* (Chl *awz*)  
 103. *Chlorophyll axz* (Chl *axz*)  
 104. *Chlorophyll ayz* (Chl *ayz*)  
 105. *Chlorophyll ayz* (Chl *ayz*)  
 106. *Chlorophyll azz* (Chl *azz*)  
 107. *Chlorophyll azaa* (Chl *aza*)  
 108. *Chlorophyll abz* (Chl *abz*)  
 109. *Chlorophyll acz* (Chl *acz*)  
 110. *Chlorophyll adz* (Chl *adz*)  
 111. *Chlorophyll aez* (Chl *aez*)  
 112. *Chlorophyll afz* (Chl *afz*)  
 113. *Chlorophyll agz* (Chl *agz*)  
 114. *Chlorophyll ahz* (Chl *ahz*)  
 115. *Chlorophyll aiz* (Chl *aiz*)  
 116. *Chlorophyll ajz* (Chl *ajz*)  
 117. *Chlorophyll akz* (Chl *akz*)  
 118. *Chlorophyll alz* (Chl *alz*)  
 119. *Chlorophyll amz* (Chl *amz*)  
 120. *Chlorophyll anz* (Chl *anz*)  
 121. *Chlorophyll aoz* (Chl *aoz*)  
 122. *Chlorophyll apz* (Chl *apz*)  
 123. *Chlorophyll aqz* (Chl *aqz*)  
 124. *Chlorophyll arz* (Chl *arz*)  
 125. *Chlorophyll asz* (Chl *asz*)  
 126. *Chlorophyll atz* (Chl *atz*)  
 127. *Chlorophyll auz* (Chl *auz*)  
 128. *Chlorophyll avz* (Chl *avz*)  
 129. *Chlorophyll awz* (Chl *awz*)  
 130. *Chlorophyll axz* (Chl *axz*)  
 131. *Chlorophyll ayz* (Chl *ayz*)  
 132. *Chlorophyll ayz* (Chl *ayz*)  
 133.

Quercus Michx.

Query Match	Best Local Similarity	Prod No	Score	JB 24	Length
	92.58	164-150	1276.8	24	2473

Best Local Similarity	F100: NO. 1.0E-150;	Index
Matches 1300: Conservative	0: Mismatch	36:

Matches	Conservative	Mismatches	GC	GC	GC
1365	9	77	36	36	36

THE UNIVERSITY OF CHICAGO

TATGGGCGCCAGCAGCAATCTGTGTCGGGAAGA(CTGGAC)AAGTGGGAGAAAGTA(GGTGG)

100

ATGGGCGCCAGCATCTGGGCAAGCTTGAACCTGGAGGCTTC

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

51 CAGGGCCGAGGACCTACATGCTGAAGCACCTGGTGGCAGCCCGAG

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

61 CTGGCCCCGGCGGCAGAAAGTGTACATGATGAAGCACCTGGTGGGCCAGCCGAG 120

[illegible]

121 CTGGAGGGCTTCGCCCCCGGCTTGGTGGAGACCGCCGAGGGCTGCAAGCAGATC 180

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

121 CTGGAGAAGTTCGCCCTGAACCCGGCCCTGCTGGAGACCAAGGAGGGCTGCCAAGCAATC 180

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525



QY 1498 CTCAGCCAGTAA 1503  
 DB 1468 CTCAGCCAGTAA 1473

RESULT 11  
 ABL39953  
 ID ABL39953 standard; DNA: 1479 BP.  
 AC ABL39953:  
 XX  
 DI 15-MAY-2002 (first entry)  
 XX  
 DE Synthetic Gag polynucleotide sequence SEQ ID NO:3.  
 XX  
 KW Human immunodeficiency virus type C; antigenic HIV type C protein;  
 immunogenic; immunisation: gag; pol: vif; vpr: tat; rev; vpu; env; nef;  
 KW immunostimulant; gene therapy: gene: ds.  
 XX  
 OS Human immunodeficiency virus type C.  
 OS Synthetic.  
 XX  
 PN WO200204493-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 PF 05-JUL-2001; 2001WO-US21241.  
 XX  
 PR 05-JUL-2000; 2000US-0610313.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 PA (UYST-) UNIV STELLERBOSCH.  
 XX  
 PI Zur Megede J. Barnett SW. Engelbrecht S. Van Rensburg EJ;  
 XX  
 DR WPI: 2002-154920/20.  
 XX  
 XX New polynucleotides encoding antigenic HIV type C polypeptides, useful  
 PT in applications including DNA immunization or generation of packaging  
 PT cell lines, particularly in gene therapy .  
 XX  
 XX Example 1; Fig 1; 23pp; English.  
 XX  
 CC The present invention describes expression cassettes comprising a  
 CC polynucleotide sequence encoding a polypeptide comprising immunogenic  
 CC HIV type C polypeptides. The expression cassettes comprise any of the  
 CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or  
 CC Nef (i). (i) have immunostimulant activity and can be used in gene  
 CC therapy. The HIV type C polynucleotides are useful in applications  
 CC including DNA immunisation, generation of packaging cell lines, and  
 CC production of HIV type C proteins. The polynucleotides are particularly  
 CC useful in gene therapy and DNA immunisation applications. ABL39942 to  
 CC ABL40054 and ABL06204 to ABL06215 represent sequences used in the  
 CC exemplification of the present invention.  
 XX  
 XX Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 other;  
 SQ

Query Match 84.6%; Score 1276.8; DB 24; Length 1479;  
 Best Local Similarity 92.5%; Pred. No. 1.6e-150;  
 Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;

QY 1 ATGGGGCCGGCCAGCATCTCGCGCGGAGAGCTGGCAAGTGGGAGAGATCCGC 63  
 DB 1 ATGGGGCCGGCCAGCATCTCGCGCGGAGAGCTGGCAAGTGGGAGAGATCCGC 60  
 QY 61 CTGGCGCCGGCGGCAAGAACCTACATGCTGAAGCACTTGGTGGGCCAGCGCGAG 120  
 DB 61 CTGGCGCCGGCGGCAAGAACCTACATGATGAAGCACTTGGTGGGCCAGCGCGAG 120  
 QY 121 CTGAGAGGGTTGCCCTGAACCCCGCCCTGCTGGAGACCCCGAGGGCTGCAAGCATC 180  
 DB 121 CTGAGAGGGTTGCCCTGAACCCCGCCCTGCTGGAGACCCCGAGGGCTGCAAGCATC 180

QY 181 ATGAGCAGCTGCAGCCGCCCTGTGCAGACACGAGAGAGCTGCAGAGCCTGTACAAAC 240  
 DB 181 ATCCGCCAGCTGCAGCCGCCCTGTGCAGACACGAGAGAGCTGCAGAGCCTGTTCAAAC 240  
 QY 241 ACCGTGCGCACACCTGTACTGCGTGCACCGCGGATCGAGTGCAGACACCAAGGAGGCC 300  
 DB 241 ACCGTGCGCACACCTGTACTGCGTGCAGACAGAGATCGAGTGCAGACACCAAGGAGGCC 300  
 QY 301 CTGGACAAGATCGAGGAGGAGCAGACAAAGTCCCAAGAGACCCAGAGGAGAG 360  
 DB 301 CTGGACAAGATCGAGGAGGAGCAGACAAAGTGCACAGACAGATCCAGCAGCCGAGGCC 360  
 QY 361 GCGGAC---GCGAAGTGCAGGAGGAGCAACTACCCATCGTGCAGAACCTGCAGGCGCAGATG 417  
 DB 361 GCGGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGCGAAGACCTGCAGGCGCAGATG 420  
 QY 418 GTGCACAGCCCATCAGCCCGCCCGCCCTGAGCGCTGGTGAAGGTGATCAGAGAGAG 477  
 DB 421 GTGCACAGCCCATCAGCCCGCCCGCCCTGAGCGCTGGTGAAGGTGATCAGAGAGAG 480  
 QY 478 GCCTTCAGCCCGCAGGTGATCCCATGTTTACCGGCCCTGAGGAGAGGCCGCCACCCCGAG 537  
 DB 481 GCCTTCAGCCCGCAGGTGATCCCATGTTTACCGGCCCTGAGGAGAGGCCGCCACCCCGAG 540  
 QY 538 GACCTGAACACAGGATGTTGAACACCGTGGCGCGCACACAGCGCGCCATGCGAGATGCTGAAG 597  
 DB 541 GACCTGAACACAGGATGTTGAACACCGTGGCGCGCACACAGCGCGCCATGCGAGATGCTGAAG 600  
 QY 558 GACACCATCAAGCAGGAGCGCGCGAGTGGGAGCGGCTGCATCCCGCTGGAGCGCGGCCGCC 657  
 DB 601 GACACCATCAAGCAGGAGCGCGCGAGTGGGAGCGGCTGCATCCCGCTGGAGCGCGGCCGCC 660  
 QY 658 GTGGCTCCCGCCAGATGCGCGCACCCCGCGGAGCGAGTGGGAGCGGCTGCATCCCGCTGGAGCGCGGCCGCC 717  
 DB 651 ATGCGCCCGCGCCAGATGCGCGCACCCCGCGGAGCGGCTGCATCCCGCTGGAGCGCGGCCGCC 720  
 QY 718 CTGAGGAGCAGATCGCCTGGATGACAGCAACACCCCGCGTGGCGTGGGCGGACATCTAC 777  
 DB 721 CTGAGGAGCAGATCGCCTGGATGACAGCAACACCCCGCGTGGCGTGGGCGGACATCTAC 780  
 QY 778 AAGCGGTGGATCATCTCGGCCCTGACACAGATCGTGGGATGTACAGCCCGCTGAGCATC 837  
 DB 781 AAGCGGTGGATCATCTCGGCCCTGACACAGATCGTGGGATGTACAGCCCGCTGAGCATC 840  
 QY 838 CTGACATTCGCCCGCAGGCGCCCAAGGAGCGCTTCGCCGACTACGTGCACCGCTTCTTCAAG 897  
 DB 841 CTGACATTCAGCAGCGCGCCCAAGGAGCGCTTCGCCGACTACGTGCACCGCTTCTTCAAG 900  
 QY 998 ACCCTGGCGCGCAGCAGGCGCCAGGAGCGTGAAGAACTGGATGACCGAGACCCCTGCTG 957  
 DB 901 ACCCTGGCGCGCAGCAGGAGCAGCAGGAGCGTGAAGAACTGGATGACCGAGACCCCTGCTG 960  
 QY 958 GTGCAGAACCCCAACCCCGCAGTGCAGAGCAACCTCGCGGCTCTCGGCCCGCGCGGCGAGC 1017  
 DB 961 GTGCAGAACCCCAACCCCGCAGTGCAGAGCAACCTCGCGGCTCTCGGCCCGCGCGGCGAGC 1020  
 QY 1018 CTGAGGAGATGATGACCGCTTCAGGCGGTGGCGCGGCCCGCGCCAGGCGCCCGCTG 1077  
 DB 1021 CTGAGGAGATGATGACCGCTTCAGGCGGTGGCGCGGCCCGCGCCAGGCGCCCGCTG 1080  
 QY 1078 CTGCGCCAGCGATGACCCAGGCGCCACAGCGTGAACATCATGATGAGAGAGACACATTC 1137  
 DB 1081 CTGCGCCAGCGATGACCCAGGCGCCACAGCGTGAACATCATGATGAGAGAGACACATTC 1137  
 QY 1138 AAGGGCCCCCGCGCGCAAGCTCAAGTGCCTCACTGCGGGAAGGAGGCGCACATCGCCAG 1197  
 DB 1138 AAGGGCCCCCGCGCGCATCTGTCAGTGTCTCACTGCGGGAAGGAGGCGCACATCGCCAG 1197  
 QY 1198 AACTGCGCGCCCGCGCAAGAGAGGCTGCTGGAATGCGGGCAAGGAGGCGCACACAGATG 1257  
 DB 1198 AACTGCGCGCCCGCGCAAGAGAGGCTGCTGGAAGTGGGGCAAGGAGGCGCACACAGATG 1257  
 QY 1258 AAGGACTGCAACCCAGCGCCAGGCCAACTTCTCTGGGCAAGATCTGCGCCCGAGCCACAGGGC 1317

```
DB 1258 AAGGACTGACCGAGCGCCAGCCAACTTCCTGGGCAAGATCTGGCCAGGCACAAAGGCG 1312
QY 1318 CCGCCGCGCAACTTCTGAGAACCGCAGCGCGCGCGCCGCGCCACGCGTGGCGACGCGCG 1377
DB 1318 CCGCCGCGCAACTTCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1359
QY 1378 CCGCCGCGCGAGAGCTTCGCTTCGAGAGAGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
DB 1360 CCGCCGCGCGAGAGCTTCGCTTCGAGAGAGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 1419
QY 1438 GACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
DB 1420 G-----ACCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1467
QY 1498 CTGAGCGCGAGTAA 1509
DB 1468 CTGAGCGCGAGTAA 1479

RESULT 12
ACR03523
ID ACR03523 standard: DNA: 4419 BP.
XX
XX ACR03523:
DT 22-MAY-2003 (first entry)
XX
DE Synthetic DNA encoding immunogenic HIV peptide #6.
XX
XX Immunogenic HIV polypeptide; human immunodeficiency virus: HIV:
KW vaccine; gene therapy; packaging cell line; humoral immune response;
KW cellular immune response; gene delivery vector; DNA immunisation;
KW ds.
XX
XX Synthetic.
OS
XX
XX W02003004657-A1.
XX
XX 16-JAN-2003.
XX
XX 05-JUL-2002; 2002WO-US21421.
XX
XX 05-JUL-2001; 2001US-303192P.
XX
XX 31-AUG-2001; 2001US-316860P.
XX
XX 16-JAN-2002; 2002US-349728P.
XX
XX 16-JAN-2002; 2002US-349793P.
XX
XX 16-JAN-2002; 2002US-349871P.
XX
XX (CHIR ) CHIRON CORP.
PA
XX
XX Zur Megede J. Barnett SW. Lian Y.
PI
XX
XX WPI: 2003-221602/21.
DR
XX
XX New synthetic polynucleotides encoding antigenic HIV type B and/or type
PT C polypeptides, useful as immunogenic compositions or vaccines for
PT generating humoral or cellular immune responses against HIV in a
PT subject, especially humans .
XX
XX Example 1; Fig 11; 262pp; English.
PS
XX
XX The invention describes a synthetic polynucleotide encoding 2 or more
CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC derived from different HIV subtypes. The polynucleotide is useful for
CC immunisation, generation of packaging cell lines, or production of HIV
CC polypeptides. The polynucleotide and its encoded proteins are useful as
CC immunogenic compositions or vaccines for generating humoral or cellular
CC immune responses against HIV in a subject, or for inducing neutralising
CC antibodies against HIV. The gene delivery vector comprising the
CC polynucleotide is also useful for DNA immunisation of, or for
CC generating an immune response (e.g. a humoral or cellular immune
CC response) in, a subject such as a mammal, particularly a human. This
```

```
CC sequence encodes a human immunodeficiency virus immunogenic peptide.
XX
XX Sequence 4419 BP; 980 A; 1582 C; 1371 G; 486 T; 0 other;
Query Match 84.5%; Score 1275.2; DB 25; Length 4413;
Best Local Similarity 92.5%; Pred. No. 2e-150;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;
QY 1 ATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
DB 7 AIGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 66
QY 61 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 67 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 126
QY 121 CTGGAGGGCTTCGCGCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
DB 127 CTGGAGAGAGTTCGCGCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 186
QY 181 ATGAGGAGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
DB 187 ATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 246
QY 241 ACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 247 ACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306
QY 301 CTGACACAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 307 CTGACACAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 366
QY 351 GCGGAC---GGCAAGGTTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 417
DB 357 GCGGACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 426
QY 418 GTGCACACAGGCGATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477
DB 427 GTGCACACAGGCGATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486
QY 478 GCGTTACGCGCGCGGAGTGATCCCGATGTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
DB 487 GCGTTACGCGCGCGGAGTGATCCCGATGTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
QY 538 GACCTGAACAGCATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
DB 547 GACCTGAACAGCATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606
QY 598 GACACCATCAAGGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657
DB 607 GACACCATCAAGGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 666
QY 658 GTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717
DB 667 ATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 726
QY 718 CTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 777
DB 727 CTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 786
QY 778 AAGCGGTGATCATCTCTGGGCTGAAACAAAGATGTTGCGGATGTACAGCCCCGTGAGCATC 837
DB 787 AAGCGGTGATCATCTCTGGGCTGAAACAAAGATGTTGCGGATGTACAGCCCCGTGAGCATC 846
QY 838 CTGGACATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897
DB 847 CTGGACATCAACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 906
QY 898 ACCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 957
DB 907 ACCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 966
QY 958 GTGCACACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1017
```

DB 967 GTGAGAGCGGACCGGACCTGCGAGAGCATCCGGCGCTCTGCGCGCGCGCGGAGC 1026  
QY 1018 CTGAGGAGATGATCAGCGGCTTGCAGAGCGGTGGCGGCGCGCGCGGAGGCGGCTG 1077  
DB 1027 CTGAGGAGATGATCAGCGGCTTGCAGAGCGGTGGCGGCGCGCGGAGGCGGCTG 1086  
QY 1078 CTGGCGGAGCGATGAGTCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1137  
DB 1087 CTGGCGGAGCGATGAGTCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1143  
QY 1138 AAGGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1197  
DB 1144 AAGGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1203  
QY 1198 AACTGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1257  
DB 1204 AACTGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1263  
QY 1258 AAGGACTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1317  
DB 1264 AAGGACTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1323  
QY 1318 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1377  
DB 1324 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1365  
QY 1378 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1437  
DB 1366 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1445  
QY 1438 GAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1497  
DB 1426 G-----ACGGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1473  
QY 1498 CTGAGCGGCGGTA 1509  
DB 1474 CTGAGCGGCGGTA 1485

RESULT 13  
ACA03522 standard; DNA: 4716 BP.

ACA03522:  
22-MAY-2003 (first entry)  
Synthetic DNA encoding immunogenic HIV peptide #5.  
Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;  
vaccine; gene therapy; packaging cell line; humoral immune response;  
cellular immune response; gene delivery vector; DNA immunisation;  
ds.  
Synthetic.  
WO2003004657-A1.  
16-JAN-2003.  
05-JUL-2002; 2002WO-US21421.  
05-JUL-2001; 2001US-303192P.  
31-AUG-2001; 2001US-316860P.  
16-JAN-2002; 2002US-349728P.  
16-JAN-2002; 2002US-349733P.  
16-JAN-2002; 2002US-349871P.  
(CHIR ) CHIRON CORP.  
Zur Megede J, Barnett SW, Lian Y;

WPI: 2003-221602/21.  
New synthetic polynucleotides encoding antigenic HIV type 3 and/or type 4 polypeptides, useful as immunogenic compositions or vaccines for generating humoral or cellular immune responses against HIV in a subject, especially humans.  
Example 1; Fig 10; 262pp; English.  
The invention describes a synthetic polynucleotide encoding 2 or more immunogenic HIV polypeptides, where at least 2 of the polypeptides are derived from different HIV subtypes. The polynucleotide is useful for immunisation, generation of packaging cell lines, or production of HIV polypeptides. The polynucleotide and its encoded proteins are useful as immunogenic compositions or vaccines for generating humoral or cellular immune responses against HIV in a subject, or for inducing neutralising antibodies against HIV. The gene delivery vector comprising the polynucleotide is also useful for DNA immunisation of, or for generating an immune response (e.g. a humoral or cellular immune response) in a subject such as a mammal, particularly a human. This sequence encodes a human immunodeficiency virus immunogenic peptide.  
Sequence 4716 BP; 1041 A; 1684 C; 1463 G; 528 T; 0 other;  
Query Match: 84.5%; Score 1275.2; DB 25; Length 4716;  
Host Local Similarity 92.5%; Pred. No. 2e-150;  
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;  
QY 1 ATGGGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60  
DB 7 ATGGGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 66  
QY 61 CTGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
DB 67 CTGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 126  
QY 121 CTGAGAGGCTTCGCGCTGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
DB 127 CTGAGAGGCTTCGCGCTGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 186  
QY 181 ATGAGCGAGCTGCGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240  
DB 187 ATCGCGGAGCTGCGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 246  
QY 241 ACCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
DB 247 ACCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 306  
QY 301 CTGCGAGAGATGCGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360  
DB 307 CTGCGAGAGATGCGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 366  
QY 361 GCGG 417  
DB 367 GCGG 426  
QY 418 GTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 477  
DB 427 GTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 486  
QY 478 GCCTTTCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 537  
DB 487 GCCTTTCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 546  
QY 538 GACCTGAACACGATGTTGAACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 597  
DB 547 GACCTGAACACGATGTTGAACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 606  
QY 598 GACACCATCAACGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 657  
DB 607 GACACCATCAACGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 666  
QY 658 GTGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 717



```
Db      367  GCGGCAAGGCAAGGTGAGCCAGAACACTACCCCATCTGTCAGAACCTGCGAGGCGCAGATG 426
QY      418  GTGCACAGCCCATFAGUCGCGCGCACCCCTGAAGGCGTGGTGAAGTATCATCAGAGAGAG 477
Db      427  GTGCACAGCCCATAGCCGCGCGCACCCCTGAAGGCGTGGTGAAGTATCATCAGAGAGAG 486
QY      478  GCCTTCAGCCGTCAGGTGATCCCATGTTACCGGCGCTGAGCGAGGCGCCACGCCGCGAG 537
Db      487  GCCTTCAGCCGCGAGGTGATCCCATGTTACCGGCGCTGAGCGAGGCGCCACGCCGCGAG 546
QY      538  GACCTGAACACGATGTTGAACACCGTGGCGCGCCACCGAGCGCCCATGCAGATGCTGAAG 597
Db      547  GACCTGAACACGATGTTGAACACCGTGGCGCGCCACCGAGCGCCCATGCAGATGCTGAAG 506
QY      598  GACACCATCAACGAGAGGCGCGCGAGTGGAGCGGCTGCACCCCGTGCAGCGCGCGCGCC 657
Db      607  GACACCATCAACGAGAGGCGCGCGAGTGGAGCGGCTGCACCCCGTGCAGCGCGCGCGCC 666
QY      658  GTGGCCCGCGCGAGATGCGGACCCCGCGCGAGCGGACATCGCGCGCGCGCGAGCGACC 717
Db      667  ATCGCCCGCGCGAGATGCGGAGCGCGCGCGAGCGGACATCGCGCGCGCGCGAGCGACC 726
QY      718  CTCAGGAGCAGATCGCTCGGATGACACGACACCCCGCGCTGCGCGTGGCGGACATCTAC 777
Db      727  CTCAGGAGCAGATCGCTCGGATGACACGACACCCCGCGCTGCGCGTGGCGGACATCTAC 786
QY      778  AAGCGGTGATCATCCCGCGCTGACACAGATGCTGCGGATGTACAGCGCGCTGACATC 837
Db      787  AAGCGGTGATCATCCCGCGCTGACACAGATGCTGCGGATGTACAGCGCGCTGACATC 846
QY      838  CTCGACATCGCGCAGGCGCCCAAGAGCGCTTCCGCGACTAGCTGAGCGGCTTCTCAG 897
Db      847  CTCGACATCAAGCAGGCGCGCGCAGAGCGCTTCCGCGACTAGCTGAGCGGCTTCTCAG 906
QY      898  ACCCTGCGCGCGAGCAGCGCGCGCGAGCGCTTGAAGACTGATGATGACCGAGCGCTGCT 957
Db      907  ACCCTGCGCGCGAGCAGCGCGCGCGAGCGCTTGAAGACTGATGATGACCGAGCGCTGCT 967
QY      958  GTGCAGAGCGCAACCGCTACTGCAAGACCATCTCTCGGAGCTCTCGGCGCGCGCGAGC 1017
Db      967  GTGCAGAGCGCAACCGCTACTGCAAGACCATCTCTCGGAGCTCTCGGCGCGCGCGAGC 1026
QY      1018  CTCGAGGATGATGATGACCGCTGCCAGCGCTGGCGCGCGCGCGCGCGCGCGCGCG 1077
Db      1027  CTCGAGGATGATGATGACCGCTGCCAGCGCTGGCGCGCGCGCGCGCGCGCGCGCG 1086
QY      1078  CTCGCGGAGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137
Db      1087  CTCGCGGAGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1143
QY      1138  AAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1197
Db      1144  AAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1203
QY      1198  AACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1257
Db      1204  AACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1263
QY      1258  AAGACTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1317
Db      1264  AAGACTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1323
QY      1318  CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db      1324  CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1365
QY      1378  CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db      1366  CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1425
QY      1438  GACCGCGAGCGCTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
```

```
2b      1426  G-----ACCGGAGAGCCCTGACCGAGCTGAAGAGCCCTTCGGCAAGACCC 1473
QY      1498  CTGAGCCAGTAA 1509
Db      1474  CTGAGCCAGAA 1485

RESULT 15
ACA03521
ID   ACA03521 standard; DNA; 5145 BP.
XX   ACA03521;
AC   ACA03521;
DT   22-MAY-2003 (first entry)
XX
DE   Synthetic DNA encoding immunogenic HIV peptide #4.
KW   immunogenic HIV polypeptide; human immunodeficiency virus; HIV;
KW   vaccine; gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation;
KW   ds.
OS   Synthetic.
XX   WO2003004657-A1.
PN   16-JAN-2003.
ID   05-JUL-2002: 2002MO-US2:421.
XX   05-JUL-2001: 2001US-303192P.
PR   31-AUG-2001: 2001US-316860P.
PR   16-JAN-2002: 2002US-349728P.
PR   16-JAN-2002: 2002US-349793P.
PR   16-JAN-2002: 2002US-349871P.
XX   (CHIR ) CHIRON CORP.
XX
PZ   Zur Megede J, Barnett SW, Lian Y;
WF   2003:221602/21.
XX
P:   New synthetic polynucleotides encoding antigenic HIV type B and/or type
PI   C polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a
XX   subject, especially humans .
XX
PS   Example 1; Fig 9; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for
CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptides. The polynucleotide and its encoded proteins are useful as
CC   immunogenic compositions or vaccines for generating humoral or cellular
CC   immune responses against HIV in a subject, or for inducing neutralising
CC   antibodies against HIV. The gene delivery vector comprising the
CC   polynucleotide is also useful for DNA immunisation of, or for
CC   generating an immune response (e.g. a humoral or cellular immune
CC   response) in a subject such as a mammal, particularly a human. This
CC   sequence encodes a human immunodeficiency virus immunogenic peptide.
XX
SQ   Sequence 5145 BP: 1130 A; 1842 C; 1599 G; 574 T; 0 other;

Query Match      84.4%; Score 1273.6; DB 25; Length 5145;
Best Local Similarity 92.4%; Pred. No. 3.1e-150;
Matches 1397; Conservative 0; Mismatches 79; Indels 36; Gaps 4;

QY      1  ATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db      7  ATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 66
QY      61  CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
```

```
Db 67 CTGCGCCCGCGCAAGAAAGTGTACATGATGAGCACTGTGTGGGCCACCGCGAG 126
Qy 121 CTGGAGGCTTGGCCCTGAAACCGCGCTGCTGGAGACCGCCAGAGGTGCAAGACGATG 190
Db 127 CTGGAAGTCTGCGCTGAAACCGCGCTGCTGGAGACCGCCAGAGGTGCAAGACGATG 196
Qy 181 ATGAAGCAGCTGGACGCGCGCTGACAGCGCGCACGAGGAGCTGGCAGGCTGTACAAC 240
Db 187 ATCGGCGAGCTGACCGCGCGCTGACAGCGCGCACGAGGAGCTGAAGAGCTGTTCNAC 246
Qy 241 ACCGTGCCACCGCTGTACTGCTGCAGCGCGGATCGAGTTCGGACACCAAGAGAGGCC 300
Db 247 ACCGTGCCACCGCTGTACTGCTGCAGCGCGGATCGAGTTCGGACACCAAGAGAGGCC 306
Qy 301 CTGGACAAAGTCGAGGAGGAGCAGACAACTCCACGAGAAAGACCCACGAGGTCAGAGAG 360
Db 307 CTGGACAAAGTCGAGGAGGAGCAGACAACTCCACGAGAAAGTCGAGGAGGTCAGAGAG 366
Qy 361 GCCGAC---GGCAAGTGAJCCCAAGTACGCCATGCTGCAGAACTCTGAGGGCCGAGATG 417
Db 367 GCCGACAAAGGCGAAGTGAJCCCAAGTACGCCATGCTGCAGAACTCTGAGGGCCGAGATG 425
Qy 418 GTGACACAGGCCATCAAGCCCGCGACCGCTGAAGCGGCTGAGAGTGAATCGAGGAGAG 477
Db 427 GTGACACAGGCCATCAAGCCCGCGACCGCTGAAGCGGCTGAGAGTGAATCGAGGAGAG 485
Qy 478 GCGTTACAGCGCGGAGTGAJCCCAAGTACGCCATGCTGCAGAACTCTGAGGGCCGAGATG 537
Db 487 GCGTTACAGCGCGGAGTGAJCCCAAGTACGCCATGCTGCAGAACTCTGAGGGCCGAGATG 546
Qy 538 GACCTGAACACGATGTTTCAACACGCTGGCGGCGCACAGGCCGCCATGACAGTGCCTGAAG 597
Db 547 GACCTGAACACGATGTTTCAACACGCTGGCGGCGCACAGGCCGCCATGACAGTGCCTGAAG 606
Qy 598 GACACCATCAACGAGGAGGCGCGGAGTGGGAGCGCTGCACCGCGTGCACCGCGCGGCC 657
Db 607 GACACCATCAACGAGGAGGCGCGGAGTGGGAGCGCTGCACCGCGTGCACCGCGCGGCC 666
Qy 658 GTGGCGCCCGCGCGAGTCCGCGACCGCGCGGAGCGAGTCGATCGCGCGCGCGCGCGCG 717
Db 667 ATCGCGCGCGCGCGAGTCCGCGACCGCGCGGAGCGAGTCGATCGCGCGCGCGCGCGCG 726
Qy 718 CTGACGAGGAGATGCGCTGGATGACACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 777
Db 727 CTGACGAGGAGATGCGCTGGATGACACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 786
Qy 778 AAGCGGTGGATCTCTGGCGCTGACACAGATGCTGGGATGTACAGCCCGCGTGGAGATC 837
Db 787 AAGCGGTGGATCTCTGGCGCTGACACAGATGCTGGGATGTACAGCCCGCGTGGAGATC 846
Qy 838 CTGACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897
Db 847 CTGACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 906
Qy 898 ACCCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 957
Db 907 ACCCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 966
Qy 958 GTGAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1017
Db 967 GTGAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1026
Qy 1018 CTGAGGAGATGATGACCGCTGCGAGGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1077
Db 1027 CTGAGGAGATGATGACCGCTGCGAGGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1086
Qy 1078 CTGCGCGAGGCGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137
Db 1087 CTGCGCGAGGCGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1146
Qy 1138 AAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1197
```

```
Db 1144 AAGGCGCCCGCGCGCGCATGCTCAAGTGTCTCAACTCGGCAAGAGGCGCAATATCGCGCG 1203
Qy 1198 AACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1257
Db 1204 AACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1263
Qy 1258 AAGGACTGCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1317
Db 1264 AAGGACTGCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1323
Qy 1318 CGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db 1324 CGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1365
Qy 1378 CGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db 1386 CGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1425
Qy 1439 GACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
Db 1426 G-----ACCGCGAGACCGTGAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1473
Qy 1496 CTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1509
Db 1474 CTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1485
```

Search completed: September 22, 2003. 17:48:06  
Job time : 446.317 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 17:26:03 : Search time 5707.74 Seconds  
(without alignments)  
10615.605 Million cell updates/sec

Title: US-09-475-704A-4  
Perfect score: 1509  
Sequence: 1 atggggcccccgcagcat.....gcggcccccgcagcagtaa 1509

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :					GenEmbi:					Result					Query					Description				

[illegible]



Cy	1081	GCGAGGCGATGACGCCAGGGCCAAACAGCGTGAACATCATGATGCAGAAGAGCAACTTCAG	1149
Eh	1081	GCGAGGCGATGACGCCAGGGCCAAACAGCGTGAACATCATGATGCAGAAGAGCAACTTCAG	1149
Cy	1141	GCGCCCCGGCGCAACGTCTAGTGCTTTCACTGGGCAAGAGGGCCACATGCCCAAGAAC	1200
Dh	1141	GCGCCCCGGCGCAACGTCTAGTGCTTTCACTGGGCAAGAGGGCCACATGCCCAAGAAC	1200
Cy	1201	TGCGCGCGCCGCCCGCAAGAAGGGCTCTGTGAAGTGGCGAAGAGGGCCACACAGATGAAG	1260
Dh	1201	TGCGCGCGCCGCCCGCAAGAAGGGCTCTGTGAAGTGGCGAAGAGGGCCACACAGATGAAG	1260
Cy	1261	GACTGCACCGAGCGCCAGGCCAACTTCTCTGGCGAAGATCTGCCCAGGCACACAAGGGCGC	1320
Dh	1261	GACTGCACCGAGCGCCAGGCCAACTTCTCTGGCGAAGATCTGCCCAGGCACACAAGGGCGC	1320
Cy	1321	CCGCGCAACTTCTTCGCGAGAACCCGAGCGCGCGCCGCCACCACCGTGCCTCCACCGCCCC	1380
Dh	1321	CCGCGCAACTTCTTCGCGAGAACCCGAGCGCGCGCCGCCACCACCGTGCCTCCACCGCCCC	1380
Cy	1381	CCGCGCGAGAGCTTCCGCTTCGAGAGAGCACCCCGCCGCCCCCAAGCAGGCCAAGGAC	1440
Dh	1381	CCGCGCGAGAGCTTCCGCTTCGAGAGAGCACCCCGCCGCCCCCAAGCAGGCCAAGGAC	1440
Cy	1441	GCGAGGCCCTACCGCGAGGCCCTTGACCGCCCTGCGCAGCGCTGTGTGGCAGCGGCCCGCTG	1500
Dh	1441	GCGAGGCCCTACCGCGAGGCCCTTGACCGCCCTGCGCAGCGCTGTGTGGCAGCGGCCCGCTG	1500
Cy	1501	AGCCAGTAA 1509	
Dh	1501	AGCCAGTAA 1509	
RESULT 3			
AX468548			
LOCUS	AX468548		
DEFINITION	Sequence 68 from Patent WO0226209.	509 bp	DNA linear PAT 16-JUL-2002
ACCESSION	AX468548		
VERSION	AX468548.1	GI:21901378	
KEYWORDS			
SOURCE	Human immunodeficiency virus 1 (HIV-1)		
ORGANISM	Human immunodeficiency virus 1		
REFERENCE	Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate lentivirus group.		
AUTHORS	O'Hagan,D., Otten,G., Donnelly,J.J., Pele,J.M., Barnett,S., Singh,M., Ulmer,J. and Dubensky,T.W.		
TITLE	Microparticles for delivery of the heterologous nucleic acids		
JOURNAL	Patent: WO 0226209-A 68 04-APR-2002;		
FEATURES	CHIRON CORPORATION (US)		
source	Location/Qualifiers		
	1..1509		
	/organism="Human immunodeficiency virus 1"		
	/mol_type="genomic DNA"		
BASE COUNT	320 a 552 c 467 g 158 t	12 others	
ORIGIN			
Query Match	99.7%	Score 1504.2;	DB 6; Length 1509;
Best Local Similarity	99.24;	Pred. No. 2.7e-149;	
Matches 1497;	Conservative 12;	Mismatches 0;	Indels 0; Gaps 0;
Cy	1	ATVGGCGCCCGCCGACGATCTCGCGGGGAGAACTGGACAATGGGAGAATCGCG	60
Dh	1	ATVGGCGCCCGCCGACGATCTCGCGGGGAGAACTGGACAATGGGAGAATCGCG	60
Cy	61	CTGCGCCCGCGCGCAGAGACACTACATGCTCAAGACCTTGCTGGCCAGCGCGAG	120
Dh	61	CTGCGCCCGCGCGCAGAGACACTACATGCTCAAGACCTTGCTGGCCAGCGCGAG	120
Cy	121	CTTGAGGGCTTTCGCCCTTGAACCCCGGCTCTGCTGGAGACCGCCGAGGCTTGAAGCAGATC	180

```
Db 121 CTGAGGCGCTTCGCCCTTAACCCCGCGCTCTCTGGAGACCCCGAGGGCTGCAAGCAGATC 180
Qy 181 ATGAGCAGCTGAGAGCGCGCTTCAGACCGGACCGAGAGCTGCGAGCTGTACAC 240
Db 181 ATGAGCAGCTGAGAGCGCGCTTCAGACCGGACCGAGAGCTGCGAGCTGTACAC 240
Qy 241 ACCGTGCGCACCTGTATTCGCTGCGACGCGGATCGAGCTCGGACACCAAGAGGCG 300
Db 241 ACCGTGCGCACCTGTATTCGCTGCGACGCGGATCGAGCTCGGACACCAAGAGGCG 300
Qy 301 CTGACAAAGTTCGAGGAGGAGACAGTCCAGCAGCAAGACCCAGAGGCGCAAGAG 360
Db 301 CTGACAAAGTTCGAGGAGGAGACAGTCCAGCAGCAAGACCCAGAGGCGCAAGAG 360
Qy 361 GCGGAGGCGAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCGAGGCGCAGATG 420
Db 361 GCGGAGGCGAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCGAGGCGCAGATG 420
Qy 421 CACGAGGCGATCGAGGCGCGCCAGCCCTGAACGCGCTGGGTGAAGGTGATCGAGGAGAGGCG 480
Db 421 CACGAGGCGATCGAGGCGCGCCAGCCCTGAACGCGCTGGGTGAAGGTGATCGAGGAGAGGCG 480
Qy 541 CTGAACACGATGTGAACACGCTGGGCGGCGACAGCGCGCCATGCGAGATGCTGAAGGAC 600
Db 541 CTGAACACGATGTGAACACGCTGGGCGGCGACAGCGCGCCATGCGAGATGCTGAAGGAC 600
Qy 601 ACCATCAACGAGGAGCGCGGAGTGGAGCGGCTGCACCCGCTGAGGCGCGCGCGCTG 660
Db 601 ACCATCAACGAGGAGCGCGGAGTGGAGCGGCTGCACCCGCTGAGGCGCGCGCGCTG 660
Qy 721 CAGGAGCAGATCGCTGGATGACAGCAACCCCGCTGCGCGCTGGCGGACATCTAGAG 780
Db 721 CAGGAGCAGATCGCTGGATGACAGCAACCCCGCTGCGCGCTGGCGGACATCTAGAG 780
Qy 781 CGGTGGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCGGTGAGCATCTG 840
Db 781 CGSTGGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCGGTGAGCATCTG 840
Qy 841 GACATCGCGCAGGCGCCCAAGGAGCCCTTCGGGACTACCTGGAGCCGCTTCTCAAGNC 900
Db 841 GACATCGCGCAGGCGCCCAAGGAGCCCTTCGGGACTACCTGGAGCCGCTTCTCAAGNC 900
Qy 901 CTGCGCGCGAGCAGCGACCCAGGAGCTGAAGACTGGATGACCGAGACCTGTGCTGGTG 960
Db 901 CTGCGCGCGAGCAGCGACCCAGGAGCTGAAGACTGGATGACCGAGACCTGTGCTGGTG 960
Qy 961 CAGAACCGCAACCCGACTGCAAGACCATCTCGCGGCTCTCGGCGCGCGGCGGACCCCTG 1020
Db 961 CAGAACCGCAACCCGACTGCAAGACCATCTCGCGGCTCTCGGCGCGCGGCGGACCCCTG 1020
Qy 1021 CAGGAGATGATGACCGCTTGCAGGGGCTGGGCGGCGCGGCGCACAGGCGCGCGCTG 1080
Db 1021 CAGGAGATGATGACCGCTTGCAGGGGCTGGGCGGCGCGGCGCACAGGCGCGCGCTG 1080
Qy 1081 CCGGAGGCGATGAGCCAGGCGCAAGCGGTGAACATCATATGATGAGAGAGACATCTCAAG 1140
Db 1081 CCGGAGGCGATGAGCCAGGCGCAAGCGGTGAACATCATATGATGAGAGAGACATCTCAAG 1140
Qy 1141 GCGCGCGCGGCGCAAGCTCAGTGGTCACTGGGCAAGGAGGCGCCACATCGCCAGAGAC 1200
Db 1141 GCGCGCGCGGCGCAAGCTCAGTGGTCACTGGGCAAGGAGGCGCCACATCGCCAGAGAC 1200
Qy 1201 TGGCGCGCGCGCGCAAGAGGCGCTGCTGGAGGTGGCGGAGGAGGCGCCACCAAGAG 1260
Db 1201 TGGCGCGCGCGCGCAAGAGGCGCTGCTGGAGGTGGCGGAGGAGGCGCCACCAAGAG 1260
```

```
Qy 1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCACGACCAAGGGCGCG 1320
Db 1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCACGACCAAGGGCGCG 1320
Qy 1321 CCGGCAACTTCCTGCAAGACCGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1321 CCGGCAACTTCCTGCAAGACCGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Qy 1381 CCGCGGAGAGCTTCGCGTTTCGAGGAGACACCCCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 1381 CCGCGGAGAGCTTCGCGTTTCGAGGAGACACCCCGCGCGCGCGCGCGCGCGCGCGCG 1440
Qy 1441 CCGGAGCGCTTACCGGAGCGCGCTGACCGCGCTGCGGACGCTGTTCGGCAGCGCGCGCTG 1500
Db 1441 CCGGAGCGCTTACCGGAGCGCGCTGACCGCGCTGCGGACGCTGTTCGGCAGCGCGCGCTG 1500
Qy 1501 AGCCAGTAA 1509
Db 1501 AGCCAGTAA 1509

RESULT 4
AX455905
LOCUS
DEFINITION
Sequence 21 from Patent WO0204493.
ACCESSION
AX455905
VERSION
AX455905.1 GI:21714897
KEYWORDS
synthetic construct
SOURCE
synthetic construct
ORGANISM
artificial sequences.
REFERENCE
1
AUTHORS
zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE
Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL
polypeptides and uses thereof
Patent: WO 0204493-A 21 17-JAN-2002;
CHIRON CORPORATION (US) : University of Stellenbosch (ZA)
FEATURES
location/Qualifiers
source
1..1509
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic Gag coding sequence of HIV strain
AF110967"
BASE COUNT
321 a 559 c 471 g 158 t
ORIGIN
Query Match 98.7%; Score 1489.8; DB 6; Length 1509;
Best Local Similarity 99.2%; Pred. No. 8.7e-148;
Matches 1497; Conservative C; Mismatches 12; Indels 0; Gaps 0;
Qy 1 ATGGCGCGCGCGCGCGCGCATCTGGCGGCGAGAGCTGGACAACTGGGAGAGATCGCG 60
Db 1 ATGGCGCGCGCGCGCGCGCATCTGGCGGCGAGAGCTGGACAACTGGGAGAGATCGCG 60
Qy 61 CTGCGCGCGCGCGCGCGCAAGCAACTACATGCTGAAGCACCTGGTGTGGCGCAGCGCGAG 120
Db 61 CTGCGCGCGCGCGCGCGCAAGCAACTACATGCTGAAGCACCTGGTGTGGCGCAGCGCGAG 120
Qy 121 CTGGAGGCGTTTCGCGCTGACCCCGCGCTGCTGGAGACCGCGGAGGCTGCAAGCAGATC 180
Db 121 CTGGAGGCGTTTCGCGCTGACCCCGCGCTGCTGGAGACCGCGGAGGCTGCAAGCAGATC 180
Qy 181 ATGAAGCAGCTGCACCGCGCGCTGCAGACCGGACCGCAGGAGCTCGCGAGCTGTACAA 240
Db 181 ATGAAGCAGCTGCACCGCGCGCTGCAGACCGGACCGCAGGAGCTGTACAA 240
Qy 241 ACCGTGGCGCACCTGTACTGTGCTGCGACCGCGCGCATCGAGGTCCCGSACACCAAGAGGCC 300
Db 241 ACCGTGGCGCACCTGTACTGTGCTGCGACCGCGCGCATCGAGGTCCCGSACACCAAGAGGCC 300
Qy 301 CTGGACACATCGAGGAGGAGTGTCTGGAAGTGGCGGAGGAGGCGCCACCAAGAG 360
```



```

Db      481  TTCAACCCGAGGAGATCCCATTTTCAACGCCCTGAGCGAGGCGCCACCCGCCAGGAC 540
QY      541  CTGAACAGATGTTGAACACCGCTGGGGGGCCACAGGCCGCCATGAGATGCTCAAGAGAC 600
Db      541  CTGAACACCATGCTGAACACCGCTGGGGGGCCACAGGCCGCCATGAGATGCTCAAGAGAC 600
QY      601  ACCATCAAGAGAGGCGCGCGAGTGGAGCGCTGTGACCCGCTGACGCCGCGCGCGCGTG 660
Db      601  ACCATCAAGAGAGGCGCGCGAGTGGAGCGCTGTGACCCGCTGACGCCGCGCGCGCGTG 660
QY      661  GCGCGCGCGCGAGATGCGCGACCCCGGGGAGGACATCGCGGCGCGCGACGACGCTG 720
Db      661  GCGCGCGCGAGATGCGCGAGCCCGCGGAGCGAGCATCGCGGACGACGACGCTG 720
QY      721  CAGGAGAGATCGCGCTGATGACGAGACGCCCGCGCGCGCGCGCGCGCGCGGACATCAAG 780
Db      721  CAGGAGAGATCGCGCTGATGACGAGAACCGCGCGCGCGCGCGCGCGCGGAGGACATCAAG 780
QY      781  CGGTGGATCATCTGGGCTGAACAAGATGCTGGGAGTACAGCGCGCGGTGAGCATCTC 840
Db      781  CGGTGGATCATCTGGGCTGAACAAGATGCTGGGAGTACAGCGCGCGGTGAGCATCTC 840
QY      841  GACATCCGCCAGGCCCAAGGACCTTTCGGGACTACCTGGACCCCTTCTTCAAGAAC 900
Db      841  GACATCAAGCAGGCGCGCAAGGAGCGCTTCCGCGACTACGTGACCGCTTCTTCAAGAAC 900
QY      901  CTGCGCGCGGAGGAGCGCGGAGGAGTGAAGACTGGATGACCGAGACCGCTGTGGTG 960
Db      901  CTGCGCGCGGAGGAGCGCGGAGGAGTGAAGACTGGATGACCGAGACCGCTGTGGTG 960
QY      961  CAGAACGCCAACCGGAGTGAAGACCATCTGCGCGCGCGCGCGCGCGCGCGCGT 1020
Db      961  CAGAACGCCAACCGGAGTGAAGACCATCTGCGCGCGCGCGCGCGCGCGCGCGT 1020
QY      1021  GAGGAGATGATGCGGCTTSCAGGGGCGTGGCGCGCGCGCGCGCGCGCGCGCGT 1080
Db      1021  GAGGAGATGATGAGCGCGCTGCGAGGGGCGTGGCGCGCGCGCGCGCGCGCGCGT 1080
QY      1081  GCGGAGGCGATGAGCGAGCGGCAAGCGGTGAACATCATGATGCAAGAGAGCACTTCAG 1140
Db      1081  GCGGAGGCGATGAGCGAGCGGCAAGCGGCAAGCGGCAAGCGGCAAGCGGCAAG 1140
QY      1141  GCGCGCGCGCGGCAAGCGTCAAGTGTTCAGTTCGCGGCAAGGAGGCGCGCGCGCG 1200
Db      1138  GCGCAACCGCATCATCAAGTGTTCAGTTCGCGGCAAGGAGTGGCGCGCGCGCGCG 1157
QY      1201  TGCGCGCGCGCGCGCGCAAGGCGTGTGGAAGTGGCGGAGGAGGCGCGCGCGCG 1260
Db      1198  TGCGCGCGCGCGCGCGCAAGGCGTGTGGAAGTGGCGGAGGAGGCGCGCGCGCG 1257
QY      1261  GACTGCACCGAGCGCGCGCGCGCAAGTTCCTGGGCAAGATCTGGCGCGCGCGCGCG 1320
Db      1258  SACTGCACCGAGCGCGCGCGCGCGCAAGTTCCTGGGCAAGATCTGGCGCGCGCGCG 1317
QY      1321  CCGGCGCACTTCCTGCAGAACCGGAGCGGCGGCGCGCGCGCGCGCGCGCGCG 1377
Db      1318  CCGGCGCACTTCCTGCAGAACCGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCG 1377
QY      1378  CCGCGCGCGCGAGAGCTTCGCGTTCGAGGAGACCGCGCGCGCGCGCGCGCGCG 1437
Db      1378  CCGCGCGCGCGAGAGCTTCGCGTTCGAGGAGACCGCGCGCGCGCGCGCGCGCG 1437
QY      1438  GACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
Db      1438  -----CGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1482
QY      1498  CTGAGCGAGTAA 1509
Db      1483  CTGAGCGAGTAA 1494
```

RESULT 5

```

AX455983
LOCUS      1491 bp
DEFINITION Sequence 99 from Patent WO0204493.
ACCESSION AX455983
VERSION    AX455983.1 GI:21714967
KEYWORDS   synthetic construct
SOURCE      synthetic construct
ORGANISM    artificial construct
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
  JOURNAL   Patent: WO 0204493-A 99 17-JAN-2002;
  CHIRON CORPORATION (US) : University of Stellenbosch (ZA)
  LOCATION/Qualifiers
    1..1491
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"
    /note="tag_TV2_C_ZAopt"
BASE COUNT 319 a 557 c 457 g 158 t
ORIGIN
Query Match      85.8%; Score 1294.6; DB 6; Length 1491;
Best Local Similarity 92.7%; Pred. No. 2.6e-127;
Matches 1401; Conservative 0; Mismatches 87; Indels 24; Gaps 3;
QY      1  ATGGCGCGCGCGCGCGAGATCTCTGCGCGCGGAGAGCGGACAACTGGGAGAGATCCGC 60
Db      1  ATGGCGCGCGCGCGCGAGATCTCTGCGCGCGGAGAGCTGGACAAAGTGGAGAGATCCGC 60
QY      61  CTGCGCGCGCGCGCGCAAGAACACTACATGCTGAAGCACTGCTGGCGAGCGCGGAG 120
Db      61  CTGCGCGCGCGCGCGCGCAAGCACTACATGCTGAAGCACTGCTGGCGAGCGCGGAG 120
QY      121  CTGGAGGCGTTGCGCTGAACCGCGGCTGCTGGAGACCGCGGAGGCTGCAAGCAGATC 180
Db      121  CTGGAGGCGTTGCGCTGAACCGCGGCTGCTGGAGACCGCGGAGGCTGCGCGCAGATC 180
QY      181  ATGAGAGAGCTGCAGCGCGCGCTGCAGACCGGCGAGGAGTCCCGAGGAGTCCCGCTTCAAC 240
Db      181  ATCAAGAGCTGCAGCGCGCGCTGCAGACCGGCGAGGAGTCCCGAGGAGTCCCGCTTCAAC 240
QY      241  ACCGTGGCGACCGCTGACTGCTGCTGACCGCGCGCTGAGGTCCCGCGACACCAAGSAGGCC 300
Db      241  ACCGTGGCGACCGCTGACTGCTGCTGACCAAGGCGATGCGAGCTGCGCGACACCAAGSAGGCC 300
QY      301  CTGGACAGATCGAGGAGGAGCAGACAAAGTCCGACGAGAGACCCACAGGCCAAGGAG 360
Db      301  CTGGACAGATCGAGGAGGAGCAGACAAAGTCCGACGAGAGACCCACAGGCCAAGGAG 360
QY      361  GCGGACGCGCAAGGTGAGCGCAGAACTTACCCCTCGTGGCAGAACTCCAGGCGCGAGTGGT 420
Db      361  GCGGACAGAGGTGAGCGCAGAACTTACCCCTCGTGGCAGAACTCCAGGCGCGAGTGGT 420
QY      421  CACGAGCGATCAGCGCGCGCGCGCGCGCTGAACCGCTGAGCGCTGGTGAAGGTGATGAGGAGGAGGCC 480
Db      421  CACGAGCGATCAGCGCGCGCGCGCGCGCTGAACCGCTGAGCGCTGGTGAAGGTGATGAGGAGGAGGCC 480
QY      481  TTCAGCGCGCGAGGTGATCCCGCATGTTACCGCGCTGAGCGAGGCGCGCGCGCGCGAG 540
Db      481  TTCAGCGCGCGAGGTGATCCCGCATGTTACCGCGCTGAGCGAGGCGCGCGCGCGCGAG 540
QY      541  CTGAACAGCATGTTGAACACCGTGGCGGCGCACCGAGCGCGCGCGCGCGCGCGCGCGCG 600
Db      541  CTGAACAGCATGCTGAACACCGTGGCGGCGCACCGAGCGCGCGCGCGCGCGCGCGCGCG 600
QY      601  ACCATCAAGAGGAGCGCGCGCGAGTGGAGCGCGCTGCGACCGCGCTGCGAGCGCGCGCGCG 660
Db      601  ACCATCAAGAGGAGCGCGCGCGAGTGGAGCGCGCTGCGACCGCGCTGCGAGCGCGCGCGCG 660
QY      661  GCGCGCGCGCGAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
```

```
|||||
Db 561 GCGCCGCGAGATGACGAGAGCCGCGGAGGAGCATGCGCGGCGACACCGAGCGCTG 720
Qy 721 CAGAGCAGATCCCTGGATGACAGAGTAAAGCCCGCGGTGCGCGTGGCGGACATCTACAG 780
Db 721 CAGAGCAGATCCCTGGATGACAGAGTAAAGCCCGCGGTGCGCGTGGCGGACATCTACAG 780
Qy 781 CGGTGGATCATCTGGCTTGACAGAGTGTGGGATGTACAGAGCCCGTGGAGCATCTG 840
Db 781 CGGTGGATCATCTGGCTTGACAGAGTGTGGGATGTACAGAGCCCGTGGAGCATCTG 840
Qy 841 GACATCCGCGAGCGCCGAGAGGCGCTTCGCGGACTACGTGACCGCTTCCTCAAGAC 900
Db 841 GACATCAAGCAGCGCGCCGAGAGGCGCTTCGCGGACTACGTGACCGCTTCCTCAAGAC 900
Qy 901 CTGCGCGCGGAGGAGCGCCAGGAGCGTGAAGACTGGATGACCGAGACCGCTGGTGTG 960
Db 901 CTGCGCGCGGAGGAGCGCCAGGAGCGTGAAGACTGGATGACCGAGACCGCTGGTGTG 960
Qy 961 CAGAGCGCAACCGCGACTCGAGAGCATCTCGCGGCTCTCGGCGCGGCGCGCCGCTG 1020
Db 961 CAGAGCGCAACCGCGACTCGAGAGCATCTCGCGGCTCTCGGCGCGGCGCGCCGCTG 1020
Qy 1021 GAGGAGATGATACCGCTCGGAGGCGTGGGCGCGCGCGGCGCACAAAGCGCGCGTGTG 1080
Db 1021 GAGGAGATGATACCGCTCGGAGGCGTGGGCGCGCGCGGCGCACAAAGCGCGCGTGTG 1080
Qy 1081 GCGGAGCGATGAGCGAGCGCACAGCGTGAACATCATGTCAGAGAGCAACITCAAG 1140
Db 1081 GCGGAGCGATGAGCGAGCGCACAGCGTGAACATCATGTCAGAGAGCAACITCAAG 1140
Qy 1141 GCGCGCGCGGCGCAAGCTCAAGTGTCTCAACTGGGCGAAGGAGGCGCACATCGCGAAGAC 1200
Db 1141 GCGCGCGCGGCGCGCTGAGTGTCTCAACTGGGCGCGAGGCGCACATCGCGCGCAC 1200
Qy 1201 TGGCGCGCGCGCGCAAGAGCGTCTGGAAGTGTGGCAAGAGGCGCACCATGATGAAG 1260
Db 1201 TGGCGCGCGCGCGCAAGAGCGTCTGGAAGTGTGGCAAGAGGCGCACCATGATGAAG 1260
Qy 1261 GACTGCAACCGCGCGCGCAAGTCTCTGGGCAAGATCTGGCGCGCGCACCAAGGCGCG 1320
Db 1261 GACTGCAACCGCGCGCGCAAGTCTCTGGGCAAGATCTGGCGCGCGCACCAAGGCGCG 1320
Qy 1321 CCGCGCAACTTCTGAGAGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db 1321 CCGCGCAACTTCTGAGAGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Qy 1378 CCGCGCGCGCGAGCTTCGCTTCGAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db 1381 CCGCGCGCGCGAGCTTCGCTTCGAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
Qy 1438 GACCGCGAGCGCTACCGCGAGCGCGCTGACCGCGCTTCGCGAGCTGTTCGCGAGCGCG 1497
Db 1432 G-----ACCGAGCGCGCTGACCGCGCTGAGAGCGCTGTTCGCGAGCGCGCGCG 1497
Qy 1498 CTGAGCGAGTAA 1509
Db 1480 CTGAGCGAGTAA 1491
```

```
RESULT 7
AX455887
LOCUS AX455887 1479 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 3 from Patent WO0204493.
ACCESSION AX455887
VERSION AX455887.1 GI:21714880
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
```

```
polypeptides and uses thereof:
Patent: WO 0204493-A 3 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
Location/Qualifiers
1. 1479
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="synthetic gag of HIV strain AF110965"
BASE COUNT 325 a 529 c 463 g 162 t
ORIGIN
Query Match 84.8%; Score 1276.8; DB 6; Length 1479;
Best local Similarity 92.5%; Pred. No. 2e-125;
Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;
Qy 1 ATGGGCGCGCGCGGAGGAGTCTCTGCGCGCGGAGGAGTGGAGAGTGGAGAGTCCCG 60
Db 1 ATGGGCGCGCGCGGAGGAGTCTCTGCGCGCGGAGGAGTGGAGAGTGGAGAGTCCCG 60
Qy 61 CTGGCGCGCGCGGAGGAGTCTCTGCGCGCGGAGGAGTGGAGAGTGGAGAGTCCCG 120
Db 61 CTGGCGCGCGCGGAGGAGTCTCTGCGCGCGGAGGAGTGGAGAGTGGAGAGTCCCG 120
Qy 121 CTGGAGGCGTTCGCGCTGAGAGCGCGCGCTCTGCGAGAGCGCGCGAGGCTGCAAGCATC 180
Db 121 CTGGAGGCGTTCGCGCTGAGAGCGCGCGCTCTGCGAGAGCGCGCGAGGCTGCAAGCATC 180
Qy 181 ATGAGGAGTCTGAGCGCGCGCTCTGCGAGAGCGCGCGAGGAGTGGAGAGTGGAGAGT 240
Db 181 ATGAGGAGTCTGAGCGCGCGCTCTGCGAGAGCGCGCGAGGAGTGGAGAGTGGAGAGT 240
Qy 241 ACCGTGCGCGCGCTGAGGAGGAGGAGTCTGCGAGAGCGCGCGAGTGGAGAGGAGGCG 300
Db 241 ACCGTGCGCGCGCTGAGGAGGAGGAGTCTGCGAGAGCGCGCGAGTGGAGAGGAGGCG 300
Qy 301 CTGGAGAGTCTGAGGAGGAGGAGTCTGCGAGAGCGCGCGAGTGGAGAGGAGGAGGAG 360
Db 301 CTGGAGAGTCTGAGGAGGAGGAGTCTGCGAGAGCGCGCGAGTGGAGAGGAGGAGGAG 360
Qy 361 GCGGAG---GCGAGGAGGAGGAGGAGTCTGCGAGAGCGCGCGAGTGGAGAGGAGGAG 417
Db 361 GCGGAGGAGGAGGAGGAGTCTGCGAGAGCGCGCGAGTGGAGAGGAGGAGGAGGAG 417
Qy 418 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 477
Db 421 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
Qy 478 GCGTTCAGCGCGCGGAGGAGTCTGCGAGAGCGCGCGAGTGGAGAGGAGGAGGAGGAGGAG 537
Db 481 GCGTTCAGCGCGCGGAGGAGTCTGCGAGAGCGCGCGAGTGGAGAGGAGGAGGAGGAGGAG 540
Qy 548 GACGTGAACAGGAGTGTGAACAGCGTGGGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 597
Db 541 GACGTGAACAGGAGTGTGAACAGCGTGGGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Qy 598 GACACCATCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 657
Db 601 GACACCATCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Qy 658 GTGGCGCGCGCGGAGGAGTCTGCGAGAGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 717
Db 661 ATCGCGCGCGCGGAGGAGTCTGCGAGAGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
Qy 718 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 777
Db 721 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Qy 778 AAGCGGTGATCATCTGCGGCGTGAACAGAGTCTGCGGATGTACAGGCGCGCGGTGAGCATC 837
Db 781 AAGCGGTGATCATCTGCGGCGTGAACAGAGTCTGCGGATGTACAGGCGCGCGGTGAGCATC 840
Qy 838 CTGGAGATCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 897
```

```

||||| 1
841 CTGGACATCAAGCAGGCGCCACAGAGGCCCTTCGCGACTACTGTGACCGCTTCACAG 900
QY
898 ACCCTGCGCGCCGAGCAGGCGCCACCGCTAGGACGTGAAGAACTGATGACCGAGACCCCTGCTG 957
Db
901 ACCCTGCGCGCCGAGCAGGCGCCACCGCTAGGACGTGAAGAACTGATGACCGAGACCCCTGCTG 960
QY
958 GTGCAGAACCCCAACCGGACTGCAAGACCACTCTGGCGCTCTGGCGCCGCGCGCCAC 1017
Db
961 GTGCAGAACCCCAACCGGACTGCAAGACCACTCTGGCGCTCTGGCGCCGCGCGCCAC 1020
QY
1018 CTGAGAGATGATGACCGCTCTGCGAGGCGTGGGGGCGCCCGCCACAAAGGCGCGCTG 1077
Db
1021 CTGAGAGATGATGACCGCTCTGCGAGGCGTGGGGGCGCCCGCCACAAAGGCGCGCTG 1080
QY
1078 CTGCGCGAGGCGATGAGCGAGGCGACAGCGTGNACATCATGATGAGAGAGCACTTC 1137
Db
1081 CTGCGCGAGGCGATGAGCGAGGCGACAGCG---AGCGTATGATGAGAGAGCACTTC 1137
QY
1138 AAGGCGCCCGCGGCGCAAGCTCAAGTGTTCACCTGGCGGCAAGGAGGCGCACATCGGCAAG 1197
Db
1138 AAGGCGCCCGCGGCGCATCTGCTCAAGTGTTCACCTGGCGGCAAGGAGGCGCACATCGGCGCG 1197
QY
1198 AACTGCGCGCGCCCGCGCAAGAGGCTGTGGAAGTGGCGGCAAGGAGGCGTACAGATG 1257
Db
1198 AACTGCGCGCGCGCCCGCGCAAGAGGCTGTGGAAGTGGCGGCAAGGAGGCGTACAGATG 1257
QY
1258 AAGGACTGTCACCGAGCGCGCAAGTCTCTGGGCAAGTCTGGCGGAGCGCCACAGGCG 1317
Db
1258 AAGGACTGTCACCGAGCGCGCAAGTCTCTGGGCAAGTCTGGCGGAGCGCGCCACAGGCG 1317
QY
1318 CGCGCGCGCAACTCTCTGCGAGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db
1318 CGCGCGCGCAACTCTCTGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
QY
1378 CGCGCGCGCGAGCTTCGCTTCGAGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db
1360 CGCGCGCGCGAGAGCTTCGCTTCGAGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1419
QY
1438 GACCGCGAGCGCTACCGCGAGCGCGCTGACCGCGCTGCGAGCGCTGTTCGCGAGCGCGCG 1497
Db
1420 G-----ACCGCGAGCGCTACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1467
QY
1498 CTGAGCGCAGTAA 1509
Db
1468 CTGAGCGCAGTAA 1479

RESULT 8
LOCUS AX468543
DEFINITION Sequence 63 from Patent WO0226209.
ACCESSION AX468543
VERSION AX468543.1 GI:21901373
KEYWORDS
SOURCE Human immunodeficiency virus 1 (HIV-1)
ORGANISM Human immunodeficiency virus 1
VIRUSES: Retroviridae; Retroviridae; Lentivirus; Primate
lentivirus group.
REFERENCE
1
AUTHORS O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
Singh,M., Ulmer,J. and Dubensky,T.W.
TITLE Microparticles for delivery of the heterologous nucleic acids
JOURNAL Patent: WO 0226209-A 53 04-APR-2002;
CHIRON CORPORATION (US)
FEATURES
Location/Qualifiers
1..1479
/organism="Human immunodeficiency virus 1"
/mol_type="genomic DNA"
/db_xref="taxon:11676"
BASE COUNT 325 a 529 c 463 g 162 t
ORIGIN
```

```

Query Match: 84.6%; Score 1276.8; DB 6; Length 1479;
Best Local Similarity 92.5%; Pred. No. 2e-125;
Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;
QY 1 ATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db
1 ATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 61 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 CTGGAGGCGCTTGGCGCTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 CTGGAGAGGTTTGGCGCTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 181 ATGAAGCAGCTGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 ATGCCCGCAGCTGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 241 ACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 241 ACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 CTGGACAGATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Db 301 CTGGACAGATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 361 GCGGAC ---GCGAGGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 417
Db 361 GCGGACAGGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 417
QY 418 GTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477
Db 421 GTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 478 GCGCTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
Db 481 GCGCTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 538 GACCTTGAACAGCGATTTGAACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
Db 541 GACCTTGAACAGCGATTTGAACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 598 GACACCATCAAGCAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657
Db 601 GACACCATCAAGCAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
QY 658 GTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717
Db 661 ATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 718 CTGCAGGAGCAGATCGCGCTGGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 777
Db 721 CTGCAGGAGCAGATCGCGCTGGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
QY 778 AAGCGGTGATCATCTTGGCGCTGAACAGATGCTGCGCGATGTACAGCGCGCGCGCGCG 837
Db 781 AAGCGGTGATCATCTTGGCGCTGAACAGATGCTGCGCGATGTACAGCGCGCGCGCGCG 840
QY 838 CTGGACATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897
Db 841 CTGGACATCAAGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 898 ACCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 957
Db 901 ACCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
QY 958 GTGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1017
Db 961 GTGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
QY 1018 CTGGAGGAGATGATGACCGCTCTGCGAGGCGTGGGGGCGCCCGCCACAAAGGCGCGCTG 1077
```

```

Db      1021  CTGGAGAGATGATGACGGCTGCCAGGGCGTGGGGGCCAGGCAACAAGGCGCGGCTG 1080
Qy      1078  CTGGGAGGCGATGACGAGGCAACAGCGTGAACATCATGATGACAGAGGCAACATTC 1137
Db      1081  CTGGGAGGCGATGAGCGAGGCAACAGC---AGGTCATGATGCAAGAGAGAACATTC 1137
Qy      1138  AAGGGCCCCCGGGCAACGTCAGTCTCTCAAGTGTGGGCAAGAGAGGCGCATGCGAAG 1197
Db      1138  AAGGGCCCCCGGGCAATGCTCAAGTCTTCAACTGGGCAAGAGAGGCGCATGCGAAG 1197
Qy      1198  AACTGGCGCGCCCGCGGAGAGAGGCTGCTGGAUTGGGCAAGAGAGGCGCAACAGATG 1257
Db      1198  AACTGGCGCGCCCGCGGAGAGAGGCTGCTGGAUTGGGCAAGAGAGGCGCAACAGATG 1257
Qy      1258  AAGGACTGCACCGAGCGGAGGCAACTTCCTTGGGCAAGATGCGCCCAAGGCG 1317
Db      1258  AAGGACTGCACCGAGCGGAGGCAACTTCCTTGGGCAAGATGCGCCCAAGGCG 1317
Qy      1318  CGCCCGGCAACTTCCTGAGAGAGCGGAGGAGCGCGCGCGCCCAAGCTGCCGACGCG 1377
Db      1318  CGCCCGGCAACTTCCTGAGAGCGG-----CCCGAGGCGCGCGCGCC 1359
Qy      1378  CCCCGCGCGAGAGTTCGCTTCGAGGAGACACCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db      1360  CCCCGCGCGAGAGTTCGCTTCGAGGAGACACCGCGCGCGCGCGCGCGCGCGCGCG 1419
Qy      1438  GACCGGAGCGCTACCGGAGCGCGCTGACGGCGCTGCGGAGCGCTGTGGGAGCGCGCG 1497
Db      1420  G-----ACCGGAGCGCGCTGACGCGCTGAAGAGCGCTGTGGGCAACGCGCG 1467
Qy      1498  CTGAGCGAGTAA 1509
Db      1468  CTGAGCGAGTAA 1479

RESULT 9
AX468547
LOCUS      1475 bp      DNA      linear      PAT 16-JUL-2002
DEFINITION      Sequence 67 from Patent WO0226209.
ACCESSION      AX468547
VERSION      AX468547.1 GI:21901377
KEYWORDS      human immunodeficiency virus 1 (HIV 1)
SOURCE      Human immunodeficiency virus 1
ORGANISM      Viruses; Retroviral viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE      1
AUTHORS      O'Hagan,D., Otten,G., Donnelly,J.J., Peio,J.M., Barnett,S.,
            Singh,M., Ulmer,J. and Dubensky,J.W.
TITLE      Microparticles for delivery of the heterologous nucleic acids
JOURNAL      Patent: WO 0226209-A 67 04-APR-2002;
            CHIRON CORPORATION (US)
FEATURES      Location/Qualifiers
            source      1..1479
                        /organism="Human immunodeficiency virus 1"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:11676"
BASE COUNT      325 a 527 c 457 g 160 t 10 others
ORIGIN
Query Match      84.2%; Score 1271.2; dB 6; Length 1479;
Best Local Similarity 91.8%; Pred. No. 7.9e-125;
Matches 138; Conservative 10; Mismatches 78; Indels 36; Gaps 4;

Qy      1  ATGGGCGCGCGCGCGAGCATCTCTGCGGCGAGAGCTGGACAAGTGGGAGAGATCGCG 60
Db      1  ATGGGCGCGCGCGCGAGCATCTCTGCGGCGGCGAGCTGGAGCGCTGGGAGCGCATCCG 60
Qy      61  CTGGGCGCGCGCGCGAGAGCACTACATGCTGAAGCACTTGTGTGGGCGAGCGCGAG 120
Db      61  CTGGGCGCGCGCGCGAGAGTGTACATGATGAAGCACTTGTGTGGGCGAGCGCGAG 120

```

```

Qy      121  CTGGAGGCGCTTCGCCCTGAACCCCGGCGCTGCTGGAGACCGCGCGAGGCTGGAGAGATC 180
Db      121  CTGGAGAGTTCGCCCTGAACCCCGGCGCTGCTGGAGACCGCGAGGCTGGAGAGATC 180
Qy      181  ATGAAGCAGCTGGAGCGCGCGCTGGAGACCGCGCGAGGCTGGAGCGCTGTACAAAC 240
Db      181  ATCGGCGAGCTGGACCGCGCGCTGGAGACCGCGCGAGGCTGGAGCGCTGTTCAAAC 240
Qy      241  ACCTGGCCACCCTGTACTCGCTGACCGCGCGCATCGAGTTCGGGACACCAAGAGGCG 300
Db      241  ACCTGGCCACCCTGTACTCGCTGACGAGAGATCGAGTSCGGGACACCAAGAGGCG 300
Qy      301  CTGGACAGATTCGAGGAGGAGAGACAGACAGTCCACAGCAAGAGATCCAGGCGCGAGG 360
Db      301  CTGGACAGATTCGAGGAGGAGAGACAGTCCACAGCAAGAGATCCAGGCGCGAGGCG 360
Qy      361  GCGGAC---GGCAGGTGAGCCAGAACTACCCATCGTCAGAACTCGAGAGCTTCGAGG 417
Db      361  GCGGACAGGCAAGGTGAGCCAGAACTACCCATCGTCAGAACTCGAGAGCTTCGAGG 420
Qy      418  GTGACAGGCGCATGACGCGCGCGCGCGCGCTGAGCGCTGGTGAAGTGTGAGAGAG 477
Db      421  GTGACAGGCGCGCATGAGCGCGCGCGCGCGCGCTGAGCGCTGGTGAAGTGTGAGAG 480
Qy      478  GCCTTCAGCGCGCGAGGTGATCCCATGTTCAACGCGCTGAGGAGGCGCGCGCGCG 537
Db      481  GCCTTCAGCGCGCGAGGTGATCCCATGTTCAACGCGCTGAGGAGGCGCGCGCGCG 540
Qy      538  GACCTGAACAGATGTTGAACCGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
Db      541  GACCTGAACAGATGTTGAACCGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Qy      598  GACACCATCAACGAGGAGCGCGCGAGTGGGACCGCTGACCGCGCTGAGCGCGCGCG 657
Db      601  GACACCATCAACGAGGAGCGCGCGAGTGGGACCGCGCGCGCGCGCGCGCGCGCGCG 660
Qy      658  GTGGCGCGCGCGCGAGATGGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717
Db      661  ATCGCGCGCGCGCGAGATGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Qy      718  CTGACGAGGAGATCGGCTGGATGACGAGCAACCGCGCGCGCGCGCGCGCGCGCGCG 777
Db      721  CTGACGAGGAGATCGGCTGGATGACGAGCAACCGCGCGCGCGCGCGCGCGCGCGCG 780
Qy      778  AAGCGTGTGATCATCTGCGGCTTGAACAAGATCGTGGGATGTACAGCGCGCGCGCG 837
Db      781  AAGCGTGTGATCATCTGCGGCTTGAACAAGATGTGCGGATGTACAGCGCGCGCGCG 840
Qy      838  CTGGACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897
Db      841  CTGGACATCAAGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Qy      898  ACCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 957
Db      901  ACCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
Qy      959  GTGACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1017
Db      961  GTGACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Qy      1018  CTGGAGGAGATGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1077
Db      1021  CTGGAGGAGATGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Qy      1078  CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137
Db      1081  CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137
Qy      1138  AAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1197
Db      1138  AAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1197
Qy      1198  AACTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1257

```







```

RESULT 12
AF201927
LOCUS       AF201927             1503 bp      DNA          linear      SYN 16-MAR-2000
DEFINITION   Synthetic construct gag protein gene, complete cds.
ACCESSION   AF201927
VERSION      AF201927.1 GI:7248702
KEYWORDS
SOURCE       synthetic construct;
             artificial sequences.
ORGANISM     1 (bases 1 to 1509)
             zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,
             Selby,M., Otten,G.R. and Barnett,S.W.
             Increased expression and immunogenicity of sequence-modified human
             immunodeficiency virus type 1 gag gene
             J. Virol. 74 (6), 2628-2635 (2006)
MEDLINE      20148954
PUBMED       10684277
REFERENCE    2 (bases 1 to 1509)
             zur Megede,J. and Barnett,S.W.
             Direct Submission
             Submitted (04-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton.
             Emeryville, CA 94608, USA
             Location/Qualifiers
             1..1509
             /organism="synthetic construct"
             /mol_type="genomic DNA"
             /db_xref="taxon:32630"
             1..1509
             /note="HIV-1SF2 p55gag"
             /codon_start=1
             /transl_table=11
             /product="gag protein"
             /protein_id="AA043628.1"
             /db_xref="GI:7248703"
             /translation="MGAASVLSGELDKWEIKRPGKKKKYKLN:VWASRELEP
             AVNGLLETSECRDILGCLQPSLGTGSEELRSLYN:VATLYCVHQIRIDYK:KEA:E
             K:EEONSKKKQAQAAAGTGNSSOVSPYVQNLQGVHQAISPT:JANWVY
             VEEKAFSEVTFMEALSFGATPDQNLMTNMTVGHGAAMOMLKTINEEAAPDRVH
             PVHAGPIAPQMRGERSDGIAGTSTLQELGNTNNPPIPVGEIYKKW:ILGINKIV
             RLYSPSTILD:ROGPEKPDYVDRFYKTLRAEQASODVKNMTETLIVONAPDCKT
             ILKALGPAATLEEMTACQGVGGPGHKARVLAEMSOVTPAT:MMORGHFNORRTY
             KCFNCGKGGHARNCRACGQWFGREGHMDKQTERANFLGKLTWPSYKGRKPNF
             LOSRPEPTAPUEESPFREGEKITPSQOEPIDKEL:YPLISLSLFGNDPSSQ"
BASE COUNT      326 a      543 c      479 g      159 t
ORIGIN
Query Match          79.6%; Score 1201; 2B 12; Length 1509;
Best Local Similarity 86.5%; Pred. No.1.8e-117;
Matches 135; Conservative 0; Mismatches 140; Indels 36; Gaps 3;

QY      1  ATGGGGCGCGCGGCGGAGGAGTCTGCGCGCGAGAGCTGGACAAGTGGGAGGAGATCCGC 60
DB      1  ATGGGGCGCGCGCGGAGGAGTCTGAGCGCGGCGGAGCTGGACAAGTGGGAGGAGATCCGC 60

QY      61  CTGGCGCGCGCGGCGGAGGAGTCTACATGCTGAGGAGCTGCTGGTGGCGGAGCGGAG 120
DB      61  CTGGCGCGCGCGGCGGAGGAGTCTACATGCTGAGGAGCTGCTGGTGGCGGAGCGGAG 120

QY      121  CTGGAGGGCTTGGCGCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGGCTGCAAGCGATC 180
DB      121  CTGGAGGGCTTGGCGCTGAACCCCGCGCTGCTGGAGACCGCGCGAGGGCTGCGCGGAGATC 180

QY      181  ATGAGGAGCTGCGACCGCGCTGCGAGACCGCGACCGAGGAGCTGCGGAGCTGTACAA 240
DB      181  ATGAGGAGCTGCGACCGCGAGCTGCGAGACCGCGACCGAGGAGCTGCGGAGCTGTACAA 240

QY      241  ACCGTGGGCGACCTGCTACTGCTGCAACCGCGCATGAGGCTCGCGGAGACCAAGAGGC 300
DB      241  ACCGTGGGCGACCTGCTACTGCTGCAACCGCGCATGAGGCTCGCGGAGACCAAGAGGC 300

QY      301  CTGGACAAGATCGAGGAGGAGCAGAACAAAGTCCCGAGCAGAGACCCAGCGGCGCAAGGAG 360

```

```
Db 1363 CCGAGCCACCGCCCGCCCGGAGAGAGCTTCGGCTTCGGGCGAGAGAGACCAACCCG 1422
Qy 1423 AAGCAGAGCCCAAGGACCGGAGGCGCTACCGGAGGCGCCCTGAGCGCCCTGGGAGAGG 1482
Db 1423 AGCCAGAGCAGGAGCGGCGGAGGCGGAGGAGGCTGACCGGCTGACGAGGCGG 1482
Qy 1483 TTGGGAGGCGCCCGCTGAGCCAGTAA 1509
Db 1483 TTGGGAGGCGCCCGGAGGAGGAGTGGGAGGCGGCTGACCGGCTGACCGGCGCCG 2539

RESULT 13
AX427930
LOCUS AX427930 9166 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 168 from Patent WO0232943.
ACCESSION AX427930
VERSION AX427930.1 GI:21538017
KEYWORDS
SOURCE
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1 Huang, Y. and Nabel, G.J.
Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
Patent: WO 0232943-A 168 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Location/Qualifiers
source 1..9166
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/notes="plasmid pVR1012X/s containing HIV genes"
BASE COUNT 2227 a 2662 c 2513 g 1764 t
ORIGIN

Query Match 78.6% Score 1186; DB 6; length 9166;
Best Local Similarity 87.7%; Pred. No. 4.1e-116;
Matches 1326; Conservative 0; Mismatches 165; Indels 21; Gaps 2;

Qy 1 ATGGGCGCCCGCGCGAGCATCTCGCGCGGCGAGCAAGCTGGACAGGTGGAGAGATCCGC 60
Db 1880 ATGGGCGCCCGCGCGAGCATCTCGCGCGGCGGAGCTGGAGCGCTGGGAGCGCATCAAG 1939
Qy 61 CTGGCGCCCGCGCGCAAGAGCACTACATCTGAAGCACTTGCTGGGCGAGGCGGAG 120
Db 1840 CTGAAGCGCCGCGCGCAAGAGCACTACATCATGAAGCACTTGCTGGGCGAGCGGAG 1599
Qy 121 CTGGAGGCTTGCGCTGAAGCGGCGGCTGAGAGCGGCGGAGCGGCGGAGGATC 180
Db 2060 CTGGAGGCTTGCGCTGAAGCGGCGGCTGAGAGCGGCGGAGCGGCGGAGGATC 2059
Qy 181 ATGAAGCAGCTGAGCGCGCGCTGAGAGCGGCGGAGCGGCGGAGCGGCGGAGGATC 240
Db 2060 ATGAAGCAGCTGAGCGCGCGCTGAGAGCGGCGGAGCGGCGGAGCGGCGGAGGATC 2119
Qy 241 ACCGTGGCGAGCTGATCTGCTGAGCGGCGGAGCGGCGGAGCGGCGGAGGAGGCG 300
Db 2120 ACCGTGGCGAGCTGATCTGCTGAGCGGCGGAGCGGCGGAGCGGCGGAGGAGGAG 2179
Qy 301 CTGGACAAGATCGAGGAGGAGGAGCAACTCTCCAGCAAGACCCAGGAGCGGAGGAG 360
Db 2180 CTGGACAAGATCGAGGAGGAGGAGCAACTCTCCAGCAAGACCCAGGAGCGGAGGAG 2239
Qy 361 GCCGAC---GCCAAGGTGAGCCAGAACTACCCCATCTGTCGAGAACTGCGAGGCGGAGTG 417
Db 2240 GCCCAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2299
Qy 418 GTGCACACGCGCATCAGCGCGCGGAGCGGCGGAGCGGCGGAGGAGGAGGAGGAG 477
Db 2300 GTGCACACGCGCATCAGCGCGCGGAGCGGCGGAGCGGCGGAGGAGGAGGAGGAG 2359
Qy 478 GCCTTCAGCCCGCGAGGCTGATCCCGCATGTTTCACCGCCCTGAGCGAGGCGCGCCCGCAG 537
```

```
Db 2360 GCCTTCAGCCCGCGAGGCTGATCCCGCATGTTACCGCCCTGAGCGAGGCGCGCCCGCAG 2419
Qy 538 GACCTGAACAGAGTGTTGAACACCGGTGGGGGCGGACACAGCGGCCCATGAGATGCTGAAG 597
Db 2420 GACCTGAACACCATGCTGAACACCGGTGGGGGCGGACACAGCGGCCCATGAGATGCTGAAG 2479
Qy 598 GACACCATCAACGAGGAGCGCGGAGTGGGAGCGGCTGACCGCTGACCGGCGGAGGCGCGCC 657
Db 2480 GACACCATCAACGAGGAGCGCGGAGTGGGAGCGGCTGACCGCTGACCGGCGGAGGCGCG 2539
Qy 558 GTGGCGCCCGCGCGAGTGGCGGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 717
Db 2540 GTGGCGCCCGCGCGAGTGGCGGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2599
Qy 718 CTGCAGGAGCAGATGCGCTGATGACGACGACGACGACGACGACGACGACGACGACGACGAC 777
Db 2600 CTGCAGGAGCAGATGACCTGATGACGACGACGACGACGACGACGACGACGACGACGACGAC 2659
Qy 778 AAGCGGTGGATCATCTCGGCTGAGCAAGATGCTGGGATGTACAGCGCGCTGAGCATC 837
Db 2660 AAGCGGTGGATCATCTCGGCTGAGCAAGATGCTGGGATGTACAGCGCGCTGAGCATC 2719
Qy 838 CTGGACATCGCGCGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 897
Db 2720 CTGGACATCAAGCAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2779
Qy 898 ACCGTGGCGCGGAGCAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 957
Db 2780 ACCGTGGCGCGGAGCAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2839
Qy 958 GTGCAGAACGCAACCGCGGAGTGCAGACCATCTCGCGGCGCTGCGCGCGCGGCGGCGGAGC 1017
Db 2840 GTGCAGAACGCAACCGCGGAGTGCAGACCATCTCGCGGCGCTGCGCGCGCGGCGGCGGAGC 2899
Qy 1018 CTGGAGGAGATGATGACCGCTGCGAGGCGGTGGGCGCGCGCGCGCGCGCGCGCGCGCGGTG 1077
Db 2900 CTGGAGGAGATGATGACCGCTGCGAGGCGGTGGGCGCGCGCGCGCGCGCGCGCGCGCGGTG 2959
Qy 1078 CTGGCGCGGAGCGATGAGCGAGCGCAAGCTGCAAGCATCATGATGAGAGAGGAGCAACTTC 1137
Db 2960 CTGGCGCGGAGCGATGAGCGAGCGCAAGCTGCAAGCATCATGATGAGAGGAGGAGCAACTTC 3019
Qy 1138 AAGGCGCGCGCGCGCGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 1197
Db 3020 AAGGCGCGCGCGCGCGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 3079
Qy 1198 AACTGCGCGCGCGCGCGCGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTG 1257
Db 3080 AACTGCGCGCGCGCGCGCGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTG 3139
Qy 1258 AAGGATGCAAGCGCGCGCGCGCGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 1317
Db 3140 AAGGATGCAAGCGCGCGCGCGCGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 3199
Qy 1318 CGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db 3200 AGGCGCGCGGATTTCTTCTCAGAGCAG-----ACCGAGCGCGCGCGCGCGCGCGCGCG 3241
Qy 1378 CGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db 3242 CCACCGAGAGAGGAGCTTCAGGTTTGGGAGAGAGAGCAACTCCCTCTCAGAGAGGAGGAG 3301
Qy 1438 GACCGGAGGCGCGCTACCGCGAGGCGCGCTGACCGCGCTGCGAGCGCTGTCGCGAGCGCGCG 1497
Db 3302 CCGATAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3361
Qy 1498 CTGAGCGCGTAA 1509
Db 3362 TCGTCAAAATAA 3373

RESULT 14
```



JOURNAL Patent: W0 0188141-A 5 22-NOV-2001;

Geneart GmbH (DE)

Location/Qualifiers

FEATURES

Source

1. 1539

/organism: "synthetic construct"

/mol\_type: "genomic DNA"

/db\_xref: "taxon:32630"

/note= "sequence with optimized codons"

BASE COUNT 366 a 508 c 502 g 163 t

ORIGIN

Query Match 75.9%; Score 1145; DB 6; Length 1539;

Best Local Similarity 85.7%; Pred. No. 1.4e-111;

Matches 1319; Conservative 0; Mismatches 190; Indels 30; Gaps 5;

Result No.	Query			ID	Description
	Score	Match	Length		
1	1475	100.0	1479	US-09-967-464-63	Sequence 63, Appl
2	1477.4	99.9	5162	US-10-190-435-18	Sequence 18, Appl
3	1477.4	99.9	3462	US-10-190-435-16	Sequence 16, Appl
4	1477.4	99.9	4419	US-10-190-435-19	Sequence 19, Appl
5	1477.4	99.9	4483	US-10-190-435-35	Sequence 35, Appl
6	1477.4	99.9	4606	US-10-190-435-34	Sequence 34, Appl
7	1477.4	99.9	4615	US-10-190-435-36	Sequence 36, Appl
8	1477.4	99.9	4702	US-10-190-435-38	Sequence 38, Appl
9	1477.4	99.9	4716	US-10-190-435-17	Sequence 17, Appl
10	1475.8	99.8	2742	US-10-190-435-20	Sequence 20, Appl
11	1475.8	99.8	3930	US-10-190-435-9	Sequence 9, Appl
12	1475.8	99.8	3930	US-10-190-435-10	Sequence 10, Appl
13	1475.8	99.8	3930	US-10-190-435-11	Sequence 11, Appl
14	1475.8	99.8	5145	US-10-190-435-12	Sequence 12, Appl
15	1474.8	99.7	4713	US-10-190-435-59	Sequence 59, Appl
16	1473.8	99.6	5184	US-10-190-435-58	Sequence 58, Appl

```
Db 61 CTGGCGCCCGCGGCAAGAGTGTACATGATGAAGCACTGCTGTGGCGCACCGCGAG 120
QY 121 CTGAGAGATTCGGCCCTGAACCCCGGCTGCTGGAGACCGAGGGCTGCAAGCAGATC 120
Db 121 CTGAGAGATTCGGCCCTGAACCCCGGCTGCTGGAGACCGAGGGCTGCAAGCAGATC 120
QY 1261 ATCCGCCAGCTGCACCGCGCCCTGCAGACGGCAGCGAGAGCTGAAGACCTGTTCAAC 240
Db 1261 ATCCGCCAGCTGCACCGCGCCCTGCAGACGGCAGCGAGAGCTGAAGACCTGTTCAAC 240
QY 241 ACCGTGGCCACCTGTACTGGTGCAGAGAGATCGAGTCCGCGACACCAAGAGGGCC 300
Db 241 ACCGTGGCCACCTGTACTGGTGCAGAGAGATCGAGTCCGCGACACCAAGAGGGCC 300
QY 301 CTGNCACAGATCGAGGAGGACAGCAACAGTGCAGCAGAGATCCAGAGGGCGAGGCC 360
Db 301 CTGNCACAGATCGAGGAGGACAGCAACAGTGCAGCAGAGATCCAGAGGGCGAGGCC 360
QY 361 GCGGCAAGGCAAGGTGAGCGCAAGCTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420
Db 361 GCGGCAAGGCAAGGTGAGCGCAAGCTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420
QY 421 GTGCAACAGGCAATCAGCCCGCGGCAACCTGAACGCTGSGTGAAGTGTATGAGAGAG 480
Db 421 GTGCAACAGGCAATCAGCCCGCGGCAACCTGAACGCTGSGTGAAGTGTATGAGAGAG 480
QY 481 GCTTCAGCCCGGAGCTGATCCCATGTGTACCGCCCTGACCGAGGGCGCCACCGCCAG 540
Db 481 GCTTCAGCCCGGAGCTGATCCCATGTGTACCGCCCTGACCGAGGGCGCCACCGCCAG 540
QY 541 GACCTGAACAGATGTTGAACACCGTGGGCGCCACCGCCGCGCATGCAATGCTGAAG 600
Db 541 GACCTGAACAGATGTTGAACACCGTGGGCGCCACCGCCGCGCATGCAATGCTGAAG 600
QY 601 GACACCATCAAGAGAGGCGCGAGTGGACACCGTGGACCGCTGCAGCGCGGGCGCC 660
Db 601 GACACCATCAAGAGAGGCGCGAGTGGACACCGTGGACCGCTGCAGCGCGGGCGCC 660
QY 661 ATCGCCCGCGGCGAGTGGGAGCGCGCGGCGGCGAGCAGATCGCGGCGCACCAAGCAGC 720
Db 661 ATCGCCCGCGGCGAGTGGGAGCGCGCGGCGGCGAGCAGATCGCGGCGCACCAAGCAGC 720
QY 721 CTGAGAGAGCAGATCGCTGGATGACCAAGCAACCGCCGCGCATCGCGTGGGAGATGAT 780
Db 721 CTGAGAGAGCAGATCGCTGGATGACCAAGCAACCGCCGCGCATCGCGTGGGAGATGAT 780
QY 781 AAGCGTGTGATATCTTGGGCTGAACAGATGCTGGCGATGTACAGCCCGCTGAGATC 840
Db 781 AAGCGTGTGATATCTTGGGCTGAACAGATGCTGGCGATGTACAGCCCGCTGAGATC 840
QY 841 CTGGACATCAAGCAGGCGCCCAAGAGCGCTTCGCGACTACGTGGAGCGCTTCTTCAAG 900
Db 841 CTGGACATCAAGCAGGCGCCCAAGAGCGCTTCGCGACTACGTGGAGCGCTTCTTCAAG 900
QY 901 ACCGTGGCGCGGAGCAGACACCGAGAGTGAAGACTGCGATGACCGACACCTGCTG 960
Db 901 ACCGTGGCGCGGAGCAGACACCGAGAGTGAAGACTGCGATGACCGACACCTGCTG 960
QY 961 GTGCAGAACGCCAACCCGAGTGTGAAGACCATCTGCGCGCTCTCGCGCCCGCGCGCAGC 1020
Db 961 GTGCAGAACGCCAACCCGAGTGTGAAGACCATCTGCGCGCTCTCGCGCCCGCGCGCAGC 1020
QY 1021 CTGGAGGAGATGATGACCGCTGCGAGGGCTGGGCGCCCGCCAGCCACAGGCGCGCG 1080
Db 1021 CTGGAGGAGATGATGACCGCTGCGAGGGCTGGGCGCCCGCCAGCCACAGGCGCGCG 1080
QY 1081 CTGGCGGAGGCGATGAGCGAGCGCCACACCGAGCTGATGATGAGAGAGCAATTCAG 1140
Db 1081 CTGGCGGAGGCGATGAGCGAGCGCCACACCGAGCTGATGATGAGAGAGCAATTCAG 1140
QY 1141 GCGCCCGCGCGCATGCTGAAGTGTCTTCAACTGCGGCAAGGAGGGCGCCACATCGCCGCG 1200
Db 1141 GCGCCCGCGCGCATGCTGAAGTGTCTTCAACTGCGGCAAGGAGGGCGCCACATCGCCGCG 1200
```

```
QY 1201 TCCCGCGCCCGCGGCAAGAGGGCTGCTGGAGTGGCGCAAGGAGGCGCAACAGTGAAG 1260
Db 1201 TCCCGCGCCCGCGGCAAGAGGGCTGCTGGAGTGGCGCAAGGAGGCGCAACAGTGAAG 1260
QY 1261 GACTGCACCGAGCGCGGCGGCAACTTCTTGGGCAAGATCTTGGCCAGACCAAGGGCGCG 1320
Db 1261 GACTGCACCGAGCGCGGCGGCAACTTCTTGGGCAAGATCTTGGCCAGACCAAGGGCGCG 1320
QY 1321 CCCGCAACTTCTTGCAGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1380
Db 1321 CCCGCAACTTCTTGCAGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1380
QY 1381 TTGAGGAGACCAACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1440
Db 1381 TTGAGGAGACCAACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1440
QY 1441 CTGAGAGCGCTTCTTGGCAACGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1479
Db 1441 CTGAGAGCGCTTCTTGGCAACGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1479
```

## RESULT 2

```
US-10-190-435-18
; Sequence 18, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUK MEDEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGBRECHT, Susan
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10190435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 3162
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRimul_C
US-10-190-435-18
```

```
Query Match 99.9%; Score 1477.4; DB 12; Length 3162;
Best Local Similarity 99.9%; Pred. No. 5.9e-292;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGCGCGCGGCGGCAAGTCTGCGCGGCGGCAAGCTGGAGCGGCAICCGC 60
Db 7 ATGGCGCGCGCGGCGGCAAGTCTGCGCGGCGGCAAGCTGGAGCGGCAICCGC 66
QY 61 CTGCGCGCGCGGCGGCAAGTGTCTACATGATGAAGACCTGTGTGGCGCGCGGAG 120
Db 67 CTGCGCGCGCGGCGGCAAGTGTCTACATGATGAAGACCTGTGTGGCGCGGAG 126
QY 121 CTGGAAGTGTGCGCTGAACCCCGCGCTGTGGAGACCGAGGCGTGTGAAGCAGATC 180
Db 127 CTGGAAGTGTGCGCTGAACCCCGCGCTGTGGAGACCGAGGCGTGTGAAGCAGATC 186
QY 181 ATCGCGAGCTGCACCCCGCGCTGCAGACCGCGGCGGAGGCTGAAGAGCGCTGTCAAC 240
Db 187 ATCGCGAGCTGCACCCCGCGCTGCAGACCGCGGCGGAGGCTGAAGAGCGCTGTCAAC 246
QY 241 ACCGTGGCGCGGCTGTACTGCTGTGCAGCAGAGATCGAGGTCCGCGACACCAAGAGGCC 300
Db 247 ACCGTGGCGCGGCTGTACTGCTGTGCAGCAGAGATCGAGGTCCGCGACACCAAGAGGCC 306
QY 301 CTGGAACAGATCGAGGAGGAGGAGACACTGCCAGCAGACAGTCCAGCGCCGAGGCC 360
Db 301 CTGGAACAGATCGAGGAGGAGGAGACACTGCCAGCAGACAGTCCAGCGCCGAGGCC 360
```

307	DB	CTGGACAAGA	TCGAGGAGAGCAGAACAAAGTTCCTACGAGAGATCCAGACGAGCCGCGAGGCG	366
361	QY	GCGCAAGCGCAAGGTGAGCCAGAAACTACCCCATCTGTGCAGANCTCCAGGSCGCAGATG	420	
367	DB	GCGCAAGGCGCAAGGTGAGCCAGAACTACCCCATCTGTGCAGANCTCCAGGSCGCAGATG	475	
421	QY	GTGCACACAGGCATCAGCCCGCCGACACCTGACAGCCCTGGGTGAAGTGTATGTAGAGAGAG	480	
427	DB	GTGCACACAGGCATCAGCCCGCCGACACCTGACAGCCCTGGGTGAAGTGTATGTAGAGAGAG	486	
481	QY	GCTTTCAGCCCGCAGGTGATCCCACTGTTCAACGGCTTGAGCGAGGGTGCACACCCGCGAG	540	
487	DB	GCTTTCAGCCCGCAGGTGATCCCACTGTTCAACGGCTTGAGCGAGGGTGCACACCCGCGAG	545	
541	QY	GACCTGAACACGATGTTTGAACACCGTGGGGGCGACAGCGCGCATCAGATGCTGAGAG	600	
547	DB	GACCTGAACACGATGTTTGAACACCGTGGGGGCGACAGCGCGCATCAGATGCTGAGAG	606	
601	QY	GACACCATCAGAGGAGGCGCCGAGTGGAGTGGACCGCTGACCCCTGACAGCGCGCGCC	660	
607	DB	GACACCATCAGAGGAGGCGCCGAGTGGAGTGGACCGCTGACCCCTGACAGCGCGCGCC	666	
661	QY	ATCGCCCGCGGCCAGATGCGCGAGCCCGCGGCGAGCGACATCGCGCGCACACGACGACCG	720	
667	DB	ATCGCCCGCGGCCAGATGCGCGAGCCCGCGGCGAGCGACATCGCGCGCACACGACGACCG	726	
721	QY	CTCGAGGAGCAGATCGGCTCTGATGACACAGCAGCCCGCCATCCCGTGGGCGACATCTAC	780	
727	DB	CTCGAGGAGCAGATCGGCTCTGATGACACAGCAGCCCGCCATCCCGTGGGCGACATCTAC	786	
781	QY	AAGCGGTGATCATCTCTGGGCTTGAAACAGATCGTSCGGATGTACAGCCCGGTGAGCAC	840	
787	DB	AAGCGGTGATCATCTCTGGGCTTGAAACAGATCGTSCGGATGTACAGCCCGGTGAGCAC	845	
841	QY	CTGGACATCAAGAGGCGCCCAAGAGCCCTTCCGGGACTAGTGGACCGCTTTCTCAAG	900	
847	DB	CTGGACATCAAGAGGCGCCCAAGAGCCCTTCCGGGACTAGTGGACCGCTTTCTCAAG	906	
901	QY	ACCTTGC CGCGCGAGCAGACACCGAGGAGTCAAGAACTGGATGACGACACCGCTGCG	960	
907	DB	ACCTTGC CGCGCGAGCAGACACCGAGGAGTCAAGAACTGGATGACGACACCGCTGCG	966	
961	QY	GTGCAGACGCGCAACCGCGACTGTGACAGACATCTCTCGGCGCTCTCGGCGCGGCGGAC	1020	
967	DB	GTGCAGACGCGCAACCGCGACTGTGACAGACATCTCTCGGCGCTCTCGGCGCGGCGGAC	1026	
1021	QY	CTGGAGGAGATGATCACCGCTGCGAGGCGTGGCGGCGCGACAGCGCGGCGGCGG	1080	
1027	DB	CTGGAGGAGATGATCACCGCTGCGAGGCGTGGCGGCGCGACAGCGCGGCGGCGG	1086	
1081	QY	CTGGCGGAGGCGATGAGCGCGAGGCCAACACAGCGTGTATGATGAGAGAGCAACTTCAAG	1140	
1087	DB	CTGGCGGAGGCGATGAGCGCGAGGCCAACACAGCGTGTATGATGAGAGAGCAACTTCAAG	1146	
1141	QY	GCGCCCGCGGCGCATCTGTCAAGTGGTTCAACTGCGCGCAGAGAGGGCGCATGTGCGCGAAC	1200	
1147	DB	GCGCCCGCGGCGCATCTGTCAAGTGGTTCAACTGCGCGCAGAGAGGGCGCATGTGCGCGAAC	1206	
1201	QY	TCCCGCGCGCGCGCCACAGAGGGCTGTGGAACTGGGGCAAGAGGGCGACACAGATCAAG	1260	
1207	DB	TCCCGCGCGCGCGCCACAGAGGGCTGTGGAACTGGGGCAAGAGGGCGACACAGATCAAG	1266	
1261	QY	GACTTGCACGAGCGCCAGCGCAACTTCTCTGGGCAAGATCTGCGCCAGGTCACAGGCGCG	1320	
1267	DB	GACTTGCACGAGCGCCAGCGCAACTTCTCTGGGCAAGATCTGCGCCAGGTCACAGGCGCG	1326	
1321	QY	CCCGGCAACTTCTCTGACAGCGCGCCGAGCCACCGCGCCCGCGCGGAGAGGCTCTGCG	1380	
1327	DB	CCCGGCAACTTCTCTGACAGCGCGCCGAGCCACCGCGCCCGCGCGGAGAGGCTCTGCG	1386	
1381	QY	TTGAGGAGACACACCCCGCGCCAGACGACGAGAGAGGACCGCGGAGACCTGACACG	1440	
1387	DB	TTGAGGAGACACACCCCGCGCCAGACGACGAGAGAGGACCGCGGAGACCTGACACG	1446	

```

Qy 1441 CTGACAGAGCCTGTTGGCAACGACGCCCTCGACGCAATAA 1479
      |||||:||||:||||:||||:||||:||||:||||:||||:
Db 2447 CTGAAGAGGCTGTCGGCAAGGAGCCCTGAGCCAGAAA 2485

RESULT 3
US-10-90-435-16
? Sequence 16, Application US/10190435
? Publication No. US20030143248A1
? GENERAL INFORMATION:
? APPLICANT: ZOR MEGEDE, Jar
? APPLICANT: BARNETT, Susan W.
? APPLICANT: LIAN, Ying
? APPLICANT: ENGELBRECHT, Susan
? APPLICANT: VAN RENSBURG, Estrelita J.
? TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGNIC HIV TYPE C
? TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
? FILE REFERENCE: PP18133.003 / 2102-18133
? CURRENT APPLICATION NUMBER: US/10/190.435
? NUMBER OF SEQ ID NOS: 319
? SOFTWARE: Patenlin Ver. 2.0
? SEQ ID NO 16
? LENGTH: 3462
? TYPE: DNA
? ORGANISM: Artificial: Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: GagProtinaSimul_C
US-10-190-435-16

Query Match 99.9%; Score 1477.4; DB 12; Length 3462;
Best Local Similarity 99.9%; Pred. No. 5.9e-292;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Qy 1 ATGGGGCGCCGCCAGCATCTCGCGGGCGGCAAGCTGACGCGCTGGGAGTGCATCCGC 60
Db 7 ATGGGGCGCCGCCAGCATCTCGCGGGCGGCAAGCTGACGCGCTGGGAGCGCATCCGC 66

Qy 61 CTGCGCCCGCGGGCAAGAATGCTACATGATGAAGCAACCTGTTGGTGGCGCCAGCCGCGAG 120
Db 67 CTGCGCCCGCGGGCAAGAATGCTACATGATGAAGCAACCTGTTGGTGGCGCCAGCCGCGAG 126

Qy 121 CTGCAGAAGCTTGGCCCTGAACCCCGGCCCTGCTGGAGACACGAGGAGGCTGCAAGCAGATC 180
Db 127 CTGCAGAAGCTTGGCCCTGAACCCCGGCCCTGCTGGAGACACGAGGAGGCTGCAAGCAGATC 186

Qy 281 ATCGCGCAGCTGACCGCCGCTGTCGACACCGGCGAGCGAGAGCTGAAGACGCTGTTCAAC 240
Db 287 ATCGCGCAGCTGACCGCCGCTGTCGACACCGGCGAGGAGGCTGAAGAGCCCTGTTCAAC 246

Qy 241 ACGGTGGCCACGCTCTACTTGGTGACAGAGAATGAGAGTCCGCGACACCAAGAGGCGC 300
Db 247 ACGTGGCCACGCTCTACTTGGTGACAGAGAATGAGAGTCCGCGACACCAAGAGGCGC 306

Qy 301 CTGCACAGATCGAGAGGACGACACAAAGTGCACGACAGAAGATCCAGCAGGCGCAGAGCC 360
Db 307 CTGCACAGATCGAGAGGACGACACAAAGTGCACGACAGAAGATCCAGCAGGCGCAGAGCC 366

Qy 361 GCGCACAGGCGCAAGTGAAGCGAGAACTACCCCATCGTGAGAACTTCGAGGCGCCAGATG 420
Db 367 GCGCACAGGCGCAAGTGAAGCGAGAACTACCCCATCGTGAGAACTTCGAGGCGCCAGATG 426

Qy 421 GTGCACAGGCGCATGAGCGCCCGCAGCCCTGAACGCGCTGGGTGAAGGTGATCGAGAGAGAG 480
Db 427 GTGCACAGGCGCATGAGCGCCCGCAGCCCTGAACGCGCTGGGTGAAGGTGATCGAGAGAGAG 486

Qy 481 GCCTTCAGCCCGAGGTGATCCCATGTTTACCGCCCTTGAGCGAGGCGCGCACCCCGCAG 540
Db 487 GCCTTCAGCCCGAGGTGATCCCATGTTTACCGCCCTTGAGCGAGGCGCGCACCCCGCAG 546

Qy 541 GACCTGAACACGATGTTGAACACCGCTGGCGGCGCCACCGAGCGGCCCATGACAGATGCTGAAG 600

```



Db 547 GACCTGAACACCATCTTCAACACCGTGGCGCCACCAGCGGCCCATSCAGATGCTGAAG 606  
QY 601 GACACCATCAAGAGAGAGCCCGCCAGTGGGACCGCGTGCACCCGCTGCACAGCGCGGCC 660  
Db 607 GACACCATCAAGAGAGAGCCCGCCAGTGGGACCGCGTGCACCCGCTGCACAGCGCGGCC 666  
QY 661 ATCGCCCGCCGCGACATGGCGGAGCCCGCGGACGACATGGCGGCGCCACCCAGCACC 720  
Db 667 ATCGCCCGCCGCGACATGGCGGAGCCCGCGGACGACATGGCGGCGCCACCCAGCACC 726  
QY 721 CTGACGAGATGATCGCTGGATGACCAAGCAACCCCGCATCCCGTGGGAGATCTAC 780  
Db 727 CTGACGAGATGATCGCTGGATGACCAAGCAACCCCGCATCCCGTGGGAGATCTAC 786  
QY 781 AAGCGTGGATCATCTGGCGCTGACAGATCTGGCGATCTACAGCCCGCTGAGATC 840  
Db 787 AAGCGTGGATCATCTGGCGCTGACAGATCTGGCGATCTGGCGATCTGGCGATCT 846  
QY 841 CTGGACATCAAGAGAGCGCCCAAGAGCGCTTCCGCGACTAGTGGACCGCTTCTCAAG 900  
Db 847 CTGGACATCAAGAGAGCGCCCAAGAGCGCTTCCGCGACTAGTGGACCGCTTCTCAAG 906  
QY 901 ACCCTGCGCGCGAGCAGACACCCAGGAGGTGAGAACTGATACCGACACCTCTCTG 960  
Db 907 ACCCTGCGCGCGAGCAGACACCCAGGAGGTGAGAACTGATGACCGACACCTCTCTG 966  
QY 961 GTGCAAGCGCAACCGCGACTGCAAGACCATCTCGCGGCTCTGGCGCGCGGCGCAGC 1020  
Db 967 GTGCAAGCGCAACCGCGACTGCAAGACCATCTCGCGGCTCTGGCGCGCGGCGCAGC 1026  
QY 1021 CTGCAAGATGATGACCGCTTGCAGAGCGGTGGCGGCTGAGCCCAAGCGCGCGCTG 1080  
Db 1027 CTGCAAGATGATGACCGCTTGCAGAGCGGTGGCGGCTGAGCCCAAGCGCGCGCTG 1086  
QY 1081 CTGGCGAGCGGATGAGCGAGCGCAACACAGCGTGTGATGCAAGAGCAACTTCAAG 1140  
Db 1087 CTGGCGAGCGGATGAGCGAGCGCAACACAGCGTGTGATGCAAGAGCAACTTCAAG 1146  
QY 1141 GCGCGCGCGCGATGATCAAGTGTTCATCTGCGCAAGAGGCGCAGTACCGCGCAAG 1200  
Db 1147 GCGCGCGCGCGATGATCAAGTGTTCATCTGCGCAAGAGGCGCAGTACCGCGCAAG 1206  
QY 1201 TCGCGCGCGCGCGAGAGCGCTGCTGCAAGTCTGCGCAAGAGGCGCAGTACCGCAAG 1260  
Db 1207 TCGCGCGCGCGCGAGAGCGCTGCTGCAAGTCTGCGCAAGAGGCGCAGTACCGCAAG 1266  
QY 1261 GACTGACAGCGCGAGCGCAACTTCTGCGCAAGATCTGCGCGAGCGCAAGAGCGCGC 1320  
Db 1267 GACTGACAGCGCGAGCGCAACTTCTGCGCAAGATCTGCGCGAGCGCAAGAGCGCGC 1326  
QY 1321 CTGCGCAACTTCTGCGAGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1380  
Db 1327 CTGCGCAACTTCTGCGAGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1386  
QY 1381 TTGCGAGAGACACCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
Db 1387 TTGCGAGAGACACCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446  
QY 1441 CTGAGAGCGCTGTGCGGCAACGACCTCCCTGAGCCAGTAA 1479  
Db 1447 CTGAGAGCGCTGTGCGGCAACGACCTCCCTGAGCCAGTAA 1485

## RESULT 4

US-10-190-435-19

; Sequence 19, Application US/10190435

; Publication No. US20030143248A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEYER, Jan

; APPLICANT: BARNETT, Susan W.

; APPLICANT: LIAN, Ying

; APPLICANT: ENGELBRECHT, Susan

; APPLICANT: VAN RENSBURG, Estrelita J.

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: P18133.003 / 2302-18133  
; CURRENT APPLICATION NUMBER: US/10190,435  
; CURRENT FILING DATE: 2002-12-30  
; NUMBER OF SEQ ID NOS: 319  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 19  
; LENGTH: 4419  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: GagRtm1a1aRevNef\_C  
US-10-190-435-19

Query Match 99.9%; Score 1477.4; DB 12; Length 4419;  
Best Local Similarity 99.9%; Pred. No. 5.8e-292;  
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGGGCGCGCGCGCCAGCATCTGCGCGCGCGCAAGCTGGAGCGCTGGGAGCGCATCCGC 60  
Db 7 ATGGGCGCGCGCGCCAGCATCTGCGCGCGCGCAAGCTGGAGCGCTGGGAGCGCATCCGC 66  
QY 61 CTGCGCGCGCGCGCGCAAGAGTCTACATGATGAGCACTTGTGTGGGCGCAGCGCGAG 120  
Db 67 CTGCGCGCGCGCGCGCAAGAGTCTACATGATGAGCACTTGTGTGGGCGCAGCGCGAG 126  
QY 121 CTGCGAGAGTCTGCGCGCTGAAGCGCGCTGCTGGAGACCGAGGCGCTGCAAGCAGATC 180  
Db 127 CTGCGAGAGTCTGCGCGCTGAAGCGCGCTGCTGGAGACCGAGGCGCTGCAAGCAGATC 186  
QY 181 ATCGCGCAGCTGCAACCGCGCGCTGAGACCGCGCAGCGAGAGCTGAAGAGCGCTGTTCAAC 240  
Db 187 ATCGCGCAGCTGCAACCGCGCGCTGAGACCGCGCAGCGAGAGCTGAAGAGCGCTGTTCAAC 246  
QY 241 ACCTGGCGCACCTGTACTTGGCTGCGACGAGAGATGCGAGTCCGCGACACCAAGAGCGCC 300  
Db 247 ACCTGGCGCACCTGTACTTGGCTGCGACGAGAGATGCGAGTCCGCGACACCAAGAGCGCC 306  
QY 301 CTGCAAGAGATCGAGGAGGAGGAGCAAGTGTGCGACGAGAGATCCAGCAGCGCGCGAGGCC 360  
Db 307 CTGCAAGAGATCGAGGAGGAGGAGCAAGTGTGCGACGAGAGATCCAGCAGCGCGCGAGGCC 366  
QY 361 GCGCAAGAGGAGGAGGAGGAGCAAGTGTGCGACGAGAGATCCAGCAGCGCGCGAGATG 420  
Db 367 GCGCAAGAGGAGGAGGAGGAGCAAGTGTGCGACGAGAGATCCAGCAGCGCGCGAGATG 426  
QY 421 GCGCAGCGCGCATGAGCGCGCGCGCGCGCGCTGAAGCGCTGAGTGAAGTGTGAGGAGAG 480  
Db 427 GCGCAGCGCGCATGAGCGCGCGCGCGCGCGCTGAAGCGCTGAGTGAAGTGTGAGGAGAG 486  
QY 481 GCGTTCAGCGCGCGAGTGTATCCCGCATGTTACCGCGCTGAGGAGGCGCGCGCGCGCGCC 540  
Db 487 GCGTTCAGCGCGCGAGTGTATCCCGCATGTTACCGCGCTGAGGAGGCGCGCGCGCGCGCC 546  
QY 541 GACTGAACACGATGTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
Db 547 GACTGAACACGATGTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606  
QY 601 GACACCATCAACGAGGAGCGCGCGCGAGTGGGACCGCTGACCGCTGCGACCGCGCGCGCG 660  
Db 607 GACACCATCAACGAGGAGCGCGCGCGAGTGGGACCGCTGACCGCTGCGACCGCGCGCGCG 666  
QY 661 ATCGCGCGCGCGCGAGATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720  
Db 667 ATCGCGCGCGCGCGAGATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 726  
QY 721 CTGCGAGGAGCAGATCCCTGGATGACCGACCAACCGCGCGCGCGCGCGCGCGCGCGCG 780  
Db 727 CTGCGAGGAGCAGATCCCTGGATGACCGACCAACCGCGCGCGCGCGCGCGCGCGCGCG 786  
QY 781 AAGCGTGGATCATCTCGCGCGCTGAAGAGAGTGTGGCGATGTACAGCGCGCTGAGCAGTAC 840

1081 CTGCGGAGGATGATGAGCCAGGCGCAATCCAGGCTGATGATGAGAGAGCAACTTCAG 1143  
1087 CTGCGGAGGATGATGAGCCAGGCGCAATCCAGGCTGATGATGAGAGAGCAACTTCAG 1146  
1141 GGGCCCCGGCGATGCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAG 1200  
1147 GGGCCCCGGCGATGCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAG 1206  
1201 TGGCGGCGCCCCCAAGAGGGGCTGCTGAAGTGGCGGAGGAGGCGCCACCAAGATCAAG 1250  
1207 TGGCGGCGCCCCCAAGAGGGGCTGCTGAAGTGGCGGAGGAGGCGCCACCAAGATCAAG 1256  
1261 GACTGCAACCTTCTGCGAGAGCGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCG 1320  
1267 GACTGCAACCTTCTGCGAGAGCGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCG 1326  
1321 CCGCGCACTTCTGCGAGAGCGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCG 1380  
1327 CCGCGCACTTCTGCGAGAGCGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCG 1386  
1381 TTGAGGAGACACCCCGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440  
1387 TTGAGGAGACACCCCGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1446  
1441 CTGAAGAGCTGTTGCGCAAGCGCCCGCTGAGCGGTTAA 1479  
1447 CTGAAGAGCTGTTGCGCAAGCGCCCGCTGAGCGGAGAA 1485

RESULT 5  
US-10-190-435-35  
; Sequence 35, Application US/13190435  
; Publication No. US20030143248A1  
; GENERAL INFORMATION:  
; APPLICANT: ZOR MEDEDE, Jan  
; APPLICANT: BARNETT, Susan W.  
; APPLICANT: LIAN, Ying  
; APPLICANT: ENGELBRECHT, Susan  
; APPLICANT: VAN RENSBURG, Estrellita J.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
; FILE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: PPI18133.003 / 2302-18-33  
; CURRENT APPLICATION NUMBER: US/10190-435  
; CURRENT FILING DATE: 2002-12-30  
; NUMBER OF SEQ ID NOS: 319  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 35  
; LENGTH: 4483  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: gp160mod.IV1.dV1v2-gagmod.BW965

Query Match 99.9%; Score 1477.4; DB 12; Length 4483;  
Best Local Similarity 99.9%; Pred. No. 5.8e-292;  
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGGGCGCGCGCGCAGCATCTCGCGGGGCGCAAGCTGGACGCTCGGAGCGGCGATCCG 50  
DB 3005 ATGGGCGCGCGCGCAGCATCTCTCGCGGGGCGCAAGCTGGACGCTCGGAGCGGCGATCCG 3064  
QY 61 CTGGCGCGCGCGCGCAGAGTGTATATATATATATATATATATATATATATATATATATAT 120  
DB 3005 CTGGCGCGCGCGCGCAGAGTGTATATATATATATATATATATATATATATATATATATAT 3124  
QY 121 CTGAGAGAGTTCGCGCTCAACCGCGCGCTCTGTGAGAGAGCGGCGGAGGCTGAGAGAGATC 180  
DB 3125 CTGAGAGAGTTCGCGCTCAACCGCGCGCTCTGTGAGAGAGCGGCGGAGGCTGAGAGAGATC 3184  
QY 181 ATGCGCCAGCTGCGCGCGCGCGCTCTGAGAGCGGCGAGGAGTGAAGAGCGCTGTTCAAC 240  
DB 3185 ATGCGCCAGCTGCGCGCGCGCGCTCTGAGAGCGGCGAGGAGTGAAGAGCGCTGTTCAAC 3244  
QY 241 ACCGTGGCGCGCGCTGCTACTGCGTGCAGAGAGATCGAGTTCGCGCGAGAGAGGCG 300  
DB 3245 ACCGTGGCGCGCGCTGCTACTGCGTGCAGAGAGATCGAGTTCGCGCGAGAGAGGCG 3304  
QY 301 CTGAGAGAGTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
DB 3305 CTGAGAGAGTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3364  
QY 361 CCGAGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
DB 3365 CCGAGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3424  
QY 421 GTGACACAGGCGCATCAGCGCGCGCGCGCGCTGAACCGCTGGGTGAAGGTGATCGAGGAGAAG 480  
DB 3425 GTGACACAGGCGCATCAGCGCGCGCGCGCGCTGAACCGCTGGGTGAAGGTGATCGAGGAGAAG 3484  
QY 481 GCTTTCAGCGCGCGAGGTGATCCCGATGTTACCGCGCTGAGCGAGGCGCGCGCGCGCGAG 540  
DB 3485 GCTTTCAGCGCGCGAGGTGATCCCGATGTTACCGCGCTGAGCGAGGCGCGCGCGCGCGAG 3544  
QY 541 GACTGACACAGTGTTCAGACCGCTGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
DB 3545 GACTGACACAGTGTTCAGACCGCTGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3604  
QY 601 GACACCATCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660  
DB 3605 GACACCATCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3664  
QY 661 ATCG 720  
DB 3665 ATCG 3724  
QY 721 CTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
DB 3725 CTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3784  
QY 781 AAGCGGTGATCTCTGCGCGCTGAACAGAGTCTGCGGAGTGTGAGAGTGTGAGAGGAGGAGG 840  
DB 3785 AAGCGGTGATCTCTGCGCGCTGAACAGAGTCTGCGGAGTGTGAGAGTGTGAGAGGAGGAGG 3844  
QY 841 CTGAGACATCAAGGAGGCG 900  
DB 3845 CTGAGACATCAAGGAGGCG 3904  
QY 901 ACCGTGCG 960  
DB 3905 ACCGTGCG 3964  
QY 961 GTGAGAGAGGCGCAACCG 1020  
DB 3965 GTGAGAGAGGCGCAACCG 4024  
QY 1021 CTGAGAGGAGTGTGATGAGCGCGCTGCGAGGCGTGGCGCGCGCGCGCGCGCGCGCGCG 1080



```
QY 1261 GACTGACCGAGCGCAGGCAACTTCCTGGGCAAGATCTGGCCACCAACAGAGGCGCG 1320
Db 4358 GACTGACCGAGCGCAGGCAACTTCCTGGGCAAGATCTGGCCACCAACAGAGGCGCG 1447
QY 1321 CCCGGCAACTTCCTGGCAGAGCGCGCCGAGCCCAACCGCCCGCCCGCGGAGAGCTTCGGC 1390
Db 4448 CCCGGCAACTTCCTGGCAGAGCGCGCCGAGCCCAACCGCCCGCCCGCGGAGAGCTTCGGC 4597
QY 1381 TTCGAGGAGACACCGCCCGGCGCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 4508 TTCGAGGAGAGACACCGCCCGGCGCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1467
QY 1441 CTGAAGAGAGCTGTTCGGCAAGCAGCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 1479
Db 4568 CTGAAGAGAGCTGTTCGGCAAGCAGCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 4606

RESULT 7
US-10-190-435-36
; Sequence 36, Application US/16190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEHDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PPI8133.003 / 2302-18433
; CURRENT APPLICATION NUMBER: US/10/190.435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 4615
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gpl60mod.Tv1.dv2-gagmod.BW965
US-10-190-435-36

Query Match 99.9% Score 1477.4; DB 12: Length 4615;
Best Local Similarity 99.9%; Pred. No. 5.8e-292;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 1ATGGGCGCGCGCGCAGCATCTCGCGCGGCGCAAGCTGGACGCTGGGAGCGGTAACGG 60
Db 3137 1ATGGGCGCGCGCGCAGCATCTCGCGCGGCGCAAGCTGGACGCTGGGAGCGGTAACGG 4196
QY 6: CTGGCGCCCGCGCGCAAGAGTGTACATGATGAGACACCTGTGTGTGGGCGCAAGCGGAG 120
Db 3137 6: CTGGCGCCCGCGCGCAAGAGTGTACATGATGAGACACCTGTGTGTGGGCGCAAGCGGAG 3256
QY 121 CTGGAGAGTTCGGCGCTGACCGCGGCTCTCTGAGACCGAGCGCTGAGAGCAATC 180
Db 3257 121 CTGGAGAGTTCGGCGCTGACCGCGGCTCTCTGAGACCGAGCGCTGAGAGCAATC 3126
QY 3257 CTGGAGAGTTCGGCGCTGACCGCGGCTCTCTGAGACCGAGCGCTGAGAGCAATC 240
Db 181 3257 CTGGAGAGTTCGGCGCTGACCGCGGCTCTCTGAGACCGAGCGCTGAGAGCAATC 4376
QY 181 ATCCGCCAGCTGCACCGCGGCTCTGAGACCGAGCGCTGAGAGCAATC 240
Db 3317 181 ATCCGCCAGCTGCACCGCGGCTCTGAGACCGAGCGCTGAGAGCAATC 4376
QY 241 ACCGTGGCAGCTGTACTGTGTGAGAGAGATCGAGTTCGCGACACCAAGAGAGGCG 300
Db 3377 241 ACCGTGGCAGCTGTACTGTGTGAGAGAGATCGAGTTCGCGACACCAAGAGAGGCG 3436
QY 301 CTGGACAGAGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 3437 301 CTGGACAGAGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3436
QY 361 GCGGACAGAGGCAAGGTGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 GCGGACAGAGGCAAGGTGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
```

```
Db 3497 GCGGACAGAGGCAAGGTGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3556
QY 421 GTGCAACAGGCAATCAGCGCCCGCGCAACCTGTAACCGCTGGGTGAAGGTGATCGAGAGAG 480
Db 3557 GTGCAACAGGCAATCAGCGCCCGCGCAACCTGTAACCGCTGGGTGAAGGTGATCGAGAGAG 3616
QY 481 SCCTTCAGCGCGCGAGGTGATCCCATGTTTCACTGCGCTGAGCGAGCGCGCGCGCGAG 540
Db 3617 GCTTTCAGCGCGCGAGGTGATCCCATGTTTCACTGCGCTGAGCGAGCGCGCGCGCGAG 3676
QY 541 GACCTGAACAGAGATGTTGAACACCGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGAG 600
Db 3677 GACCTGAACAGAGATGTTGAACACCGCTGGGCGCGCGCGCGCGCGCGCGCGCGAG 3736
QY 601 CACACCATCAACAGAGAGAGCGCGCGAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 3737 GACACCATCAACAGAGAGAGCGCGCGAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3796
QY 661 ATCGCGCGCGCGCGCGAGATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Db 3797 ATCGCGCGCGCGCGCGAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3856
QY 721 CTGCAAGGAGAGATCGCTGGATGACCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
Db 3857 CTGCAAGGAGAGATCGCTGGATGACCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3916
QY 781 AAGCGGTGGATCTCTGGGCTGTAACAGATCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
Db 3917 AAGCGGTGGATCTCTGGGCTGTAACAGATCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 3976
QY 841 CTGGACATCAAGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db 3977 CTGGACATCAAGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4036
QY 901 ACCCTGCGCGCGCGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
Db 4037 ACCCTGCGCGCGCGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4096
QY 961 GTGCAAGACCGCAACCGCGCGCTGCAAGACCATCTCTGCGCGCGCGCGCGCGCGCGCGCG 1020
Db 4097 GTGCAAGACCGCAACCGCGCGCTGCAAGACCATCTCTGCGCGCGCGCGCGCGCGCGCGCG 4156
QY 1021 CTGGAGAGAGATGATGACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Db 4157 CTGGAGAGAGATGATGACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4216
QY 1081 CTGGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Db 4217 CTGGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4276
QY 1141 GCGCGCGCGCGCGCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1200
Db 4277 GCGCGCGCGCGCGCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4336
QY 1201 TGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
Db 4337 TGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4396
QY 1261 GACTGCACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
Db 4397 GACTGCACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4456
QY 1321 CCGCGCAACTTCCTGTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 4457 CCGCGCAACTTCCTGTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4516
QY 1381 TTCGAGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 4517 TTCGAGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4576
QY 1441 CTGAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1479
Db 4577 CTGAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4615
```

## RESULT 8

US-10-190-435-38

; Sequence 38, Application US/10190435

; Publication No. US20030143248A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEDEDE, Jan

; APPLICANT: BARNETT, Susan W.

; APPLICANT: LIAN, Ying

; APPLICANT: ENGELBRECHT, Susan

; APPLICANT: VAN RENSBURG, Estrelita J.

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

; FILE REFERENCE: PP18133.003 / 2302-18133

; CURRENT APPLICATION NUMBER: US/10/190.435

; NUMBER OF SEQ ID NOS: 319

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 38

; LENGTH: 4702

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: gp160mod.TV1-gagmod.HW965

US-10-190-435-38

Query Match 99.9%; Score 1477.4; DB 12; Length 4702;

Best Local Similarity 99.9%; Pred. No. 5.8e-292;

Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCGCGCGCCAGCATCTCGCGCGCGCAAGCTGGACGCTGGAGCGGCATCCGC 60  
 DB 3224 ATGGCGCGCGCGCCAGCATCTCGCGCGCGCAAGCTGGAGCGGCATCCGC 3293  
 QY 61 CTGGCGCGCGCGCGCAAGTGTCTACATGATGAAGACCTGGTGTGGCGCGAGCGGAG 120  
 DB 3284 CTGGCGCGCGCGCGCAAGTGTCTACATGATGAAGACCTGGTGTGGCGCGAGCGGAG 3343  
 QY 121 CTGGAGAACTTCGCCCTGCAACCGCGGCTGCTGGAGACAGCGGCGCTGCAACATATC 180  
 DB 3344 CTGGAGAACTTCGCCCTGCAACCGCGGCTGCTGGAGACAGCGGCGCTGCAACATATC 3403  
 QY 181 ATGGCGCGAGCTGACCGCGCGCTGAGACCGCGCGAGAGAGAGAGAGAGAGAGAGAG 240  
 DB 3404 ATGGCGCGAGCTGACCGCGCGCTGAGACCGCGCGAGAGAGAGAGAGAGAGAGAGAG 3463  
 QY 241 ACCGTGGCCACCTGTACTGCTGCACGAGAGATGCGAGGTCCGCGACAGCAAGCGCC 300  
 DB 3464 ACCGTGGCCACCTGTACTGCTGCACGAGAGATGCGAGGTCCGCGACAGCAAGCGCC 3523  
 QY 301 CTGGCAAGATCAGGAG 360  
 DB 3524 CTGGCAAGATCAGGAG 3583  
 QY 361 GCGGCAAGGCAAGGTGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
 DB 3584 GCGGCAAGGCAAGGTGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4243  
 QY 421 GTGGCAAGGCAAGGTGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 DB 3644 GTGGCAAGGCAAGGTGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3703  
 QY 481 GCGTTGAGCGCGAGGTGATCCCATGTTACCGCGCTGAGCGAGGCGCCACCGCCGAG 540  
 DB 3704 GCGTTGAGCGCGAGGTGATCCCATGTTACCGCGCTGAGCGAGGCGCCACCGCCGAG 3763  
 QY 541 GACCTGAACACATGTTGAACACCGTGGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 DB 3764 GACCTGAACACATGTTGAACACCGTGGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAG 3823  
 QY 601 GACACCATCAACGAG 660

DB 3824 GACACCATCAACGAG 3883  
 QY 661 ATCGCGCGCGCGCGAGATGCGGAGCGCGCGCGCGCGAGAGATGCGGCGCGAGAGAGAG 720  
 DB 3884 ATCGCGCGCGCGCGAGATGCGGAGCGCGCGCGCGCGAGAGATGCGGCGCGAGAGAGAG 3943  
 QY 721 CTGAGGAGAGATGCGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
 DB 3944 CTGAGGAGAGATGCGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4003  
 QY 781 AAGCGGTGGATCATCTCGGCGCTGAAACAGATGCTGGGATGTACAGCCCGCTGAGCATC 840  
 DB 4004 AAGCGGTGGATCATCTCGGCGCTGAAACAGATGCTGGGATGTACAGCCCGCTGAGCATC 4063  
 QY 841 CTGGACATCAAGCAGGCG 900  
 DB 4064 CTGGACATCAAGCAGGCG 4123  
 QY 901 ACCGTGGCGCGCGAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
 DB 4124 ACCGTGGCGCGCGAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4183  
 QY 961 GTGAGAAAGCG 1020  
 DB 4184 GTGAGAAAGCG 4243  
 QY 1021 CTGGAGAGATGATGACCGCGCTGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080  
 DB 4244 CTGGAGAGATGATGACCGCGCTGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4303  
 QY 1081 CTGGCGCGAGCGATGAGCG 1140  
 DB 4304 CTGGCGCGAGCGATGAGCG 4363  
 QY 1141 GGCG 1200  
 DB 4364 GGCG 4423  
 QY 1201 TGCG 1260  
 DB 4424 TGCG 4483  
 QY 1261 GACTGACCGAGCG 1320  
 DB 4484 GACTGACCGAGCG 4543  
 QY 1321 CCGCGCAACTTCTGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
 DB 4544 CCGCGCAACTTCTGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4603  
 QY 1381 TTGAGGAG 1440  
 DB 4604 TTGAGGAG 4663  
 QY 1441 CTGAAGAGCGCTGTGCGCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1479  
 DB 4664 CTGAAGAGCGCTGTGCGCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4702

## RESULT 9

US-10-190-435-17

; Sequence 17, Application US/10190435

; Publication No. US20030143248A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEDEDE, Jan

; APPLICANT: BARNETT, Susan W.

; APPLICANT: LIAN, Ying

; APPLICANT: ENGELBRECHT, Susan

; APPLICANT: VAN RENSBURG, Estrelita J.

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

; FILE REFERENCE: PP18133.003 / 2302-18133





```
QY 241 ACCGTGGCCACCTGTACTGCGTGCACGAGAACTCGAGTCTGGGACACCAACCAAGAGGCGC 300
DB 247 ACCGTGGCCACCTGTACTGCGTGGCAGGAGAACTCGAGTCTGGGACACCAACCAAGAGGCGC 306
QY 301 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 307 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 366
QY 361 GCGGACAAGGCGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTCGGAGGAGGAGGAGGAGG 420
DB 367 GCGGACAAGGCGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTCGGAGGAGGAGGAGGAGG 426
QY 421 GTCGACAGGCGCATACGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
DB 427 GTCGACAGGCGCATACGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 486
QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTACCGGCGCTGAGCAGGAGGCGCCACCGCCGAGGAG 540
DB 487 GCCTTCAGCCCCGAGGTGATCCCATGTTACCGGCGCTGAGCAGGAGGCGCCACCGCCGAGGAG 546
QY 541 GACCTGACAGCATTTGACACACCTGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 547 GACCTGACAGCATTTGACACACCTGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 606
QY 601 GACACCATCAACAGGAGGCGCGGAGTGGGAGCGGCTGACCGCGCTGACCGCGGCGCGCGCGCG 660
DB 607 GACACCATCAACAGGAGGCGCGGAGTGGGAGCGGCTGACCGCGCTGACCGCGGCGCGCGCGCG 666
QY 661 ATCGCCCGCGCGGAGATGGGAGGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
DB 667 ATCGCCCGCGCGGAGATGGGAGGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 726
QY 721 CTCGAGGAGAGATCGCTGGATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 780
DB 727 CTCGAGGAGAGATCGCTGGATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 786
QY 781 AAGCGGTGATCATCTCTGGGCGCTGACCAAGATCGTGGGATGTACAGCGCGCTGACGATC 840
DB 787 AAGCGGTGATCATCTCTGGGCGCTGACCAAGATCGTGGGATGTACAGCGCGCTGACGATC 846
QY 841 CTGGACATCAAGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 847 CTGGACATCAAGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906
QY 901 ACCCTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
DB 907 ACCCTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 966
QY 961 GTGCGAGACGCCAACCCCGACTGCAAGAGCATCTCTGGGCGCTGCGGCGCGCGCGCGCGCGCG 1020
DB 967 GTGCGAGACGCCAACCCCGACTGCAAGAGCATCTCTGGGCGCTGCGGCGCGCGCGCGCGCGCG 1026
QY 1021 CTGGAGGAGATGATACCGCTGCGGAGGCGTGGGCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1027 CTGGAGGAGATGATACCGCTGCGGAGGCGTGGGCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAG 1086
QY 1081 CTGGCGGAGGAGATGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
DB 1087 CTGGCGGAGGAGATGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1146
QY 1141 GCGCGCGCGCGCATGCTCAAGTGTCTCACTCGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
DB 1147 GCGCGCGCGCGCATGCTCAAGTGTCTCACTCGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1206
QY 1201 TGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 1207 TGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1266
QY 1261 GACTGCAACGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
DB 1267 GACTGCAACGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1326
```

```
QY 1321 CCGCGCAACTTCTCTGACAGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
DB 1327 CCGCGCAACTTCTCTGACAGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1386
QY 1381 TCGAGGAGAGCACCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
DB 1387 TCGAGGAGAGCACCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1446
QY 1441 CCGAGAGAGGCTGTCTGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1479
DB 1447 CTGAAGAGGCTGTCTGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1485

RESULT 12
US-10-190-435-10
: Sequence ID, Application US/10190435
: Publication No. US20030143248A1
: GENERAL INFORMATION:
: APPLICANT: ZUR MEGEDE, Jan
: APPLICANT: BARNETT, Susan W.
: APPLICANT: LIAN, Ying
: APPLICANT: ENGELBRECHT, Susan
: APPLICANT: VAN RENSBURG, Estrellita J.
: TITLE OF INVENTION: POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C
: TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
: FILE REFERENCE: P18133.903 / 2302-18133
: CURRENT APPLICATION NUMBER: US/10/190.435
: CURRENT FILING DATE: 2002-12-30
: NUMBER OF SEQ ID NOS: 319
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 10
: LENGTH: 3930
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: GagComplPolmutAtt_C
US-10-190-435-10
```

```
Query Match 99.8% Score 1475.8 DB 12 Length 3930;
Best Local Similarity 99.9% Pred. No. 1.2e-291;
Matches 1477 Conservative 0 Mismatches 2 Indels 0 Gaps 0;
```

```
QY 1 ATGGGCGCGCGCGCGCGCGCATCTCTGCGGCGGCGCAAGCTGGACGCTGGAGCGGCGCATCCGC 60
DB 7 ATGGGCGCGCGCGCGCGCGCATCTCTGCGGCGGCGCAAGCTGGACGCTGGAGCGGCGCATCCGC 66
QY 61 CTGGCGCGCGCGCGCGCGAGAGTGTCTACATGATGAAGACCTGGTGTGGCGCGCGCGGAG 120
DB 67 CTGGCGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGACCTGGTGTGGCGCGCGCGGAG 126
QY 121 CTGGAGAGGTTTCGCGCTGAACCGCGGCTGTGGAGACCGAGCGAGGCTGCAAGCAGATC 180
DB 127 CTGGAGAGGTTTCGCGCTGAACCGCGGCTGTGGAGACCGAGCGAGGCTGCAAGCAGATC 186
QY 131 ATCGCGCGAGCTGGACCGCGCGCTGCAGACCGCGCGCGCGAGGAGGCTGAAGCGCTTTCAC 240
DB 137 ATCGCGCGAGCTGGACCGCGCGCTGCAGACCGCGCGCGAGGAGGCTGAAGCGCTTTCAC 246
QY 241 ACCGTGGCGCACCTGTACTGTGTGACGAGAGAGATCGAGGTCCGCGACACCAAGGAGGCGC 300
DB 247 ACCGTGGCGCACCTGTACTGTGTGACGAGAGAGATCGAGGTCCGCGACACCAAGGAGGCGC 306
QY 301 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 307 CTGGACAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 366
QY 361 GCGGACAAGGCGAGGTGAGCGCAGAACTACCCCATCGTGCAGAACCTGCGAGGCGCGCAGATG 420
DB 367 GCGGACAAGGCGAGGTGAGCGCAGAACTACCCCATCGTGCAGAACCTGCGAGGCGCGCAGATG 426
QY 421 GTGCAGCAGGCGCATACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 427 GTGCAGCAGGCGCATACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486
```







Db 907 ACCGTGGCGCCGAGCAGAGACACCCAGAGGTGAAGAACTGGATGACCGAGACGCTGCTG 968  
QY 961 GTGCAGAAAGCCACACCGGACTGCAAGAACCATCTGTGCGCTCTGCGCCCGCGCCAGC 1020  
Db 967 GTGCAGAAAGCCACACCGGACTGCAAGAACCATCTGTGCGCTCTGCGCCCGCGCGCAGC 1026  
QY 1021 CTGCAGAGATGATGACCGGCTGCCAGGGGTGGCGGCCCGCCAGGCCACAAAGCGCGCGTG 1080  
Db 1027 CTGCAGAGATGATGACCGGCTGCCAGGGGTGGCGGCCCGCCAGGCCACAAAGCGCGCGTG 1086  
QY 1081 CTGCAGAGAGGGAAGAGCCAGGCCAACACCAAGCGGTGATGATGAGAGAGACAACTTCAAG 1140  
Db 1087 CTGCAGAGAGGGAAGAGCCAGGCCAACCAACCAAGCGGTGATGATGAGAGAGACAACTTCAAG 1146  
QY 1141 GGGCCCGCGGCGCACTGTCAGTGTCTCAACTGGGGCAAGAGAGGCCACATCGCCCGCAAC 1200  
Db 1147 GGGCCCGCGGCGCACTGTCAGTGTCTCAACTGGGGCAAGAGAGGCCACATCGCCCGCAAC 1206  
QY 1201 TGCGCGCCCGCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGAGGCCACCAATGAAG 1260  
Db 1207 TGCGCGCCCGCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGAGGCCACCAATGAAG 1266  
QY 1261 GACTGCAGAGAGCCCGCCAGGCAACTTCTGGGCAAGATCTGGGCCAGGCCACAAAGGCGCG 1320  
Db 1267 GACTGCAGAGAGCCCGCCAGGCAACTTCTGGGCAAGATCTGGGCCAGGCCACAAAGGCGCG 1326  
QY 1321 CCGCGCAACTTCTGTCAGAGCGCGCCGAGCCACCGCCCGCCCGCCCGCGGAGGCTTCGCG 1380  
Db 1327 CCGCGCAACTTCTGTCAGAGCGCGCCGAGCCACCGCCCGCCCGCCCGCGGAGGCTTCGCG 1386  
QY 1381 TTCAGAGAGAGCCCGCCCGCCAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
Db 1387 TTCAGAGAGAGCCCGCCCGCCAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446  
QY 1441 CTGAAGAGCGCTGTCGGCAAGAGCGCCCTGAGCGAGTAA 1479  
Db 1447 CTGAAGAGCGCTGTCGGCAAGAGCGCCCTGAGCGAGAA 1485

## RESULT 15

US-10-190-435-59  
; Sequence 59, Application: US/1019435  
; Publication No.: US20030143248A1  
; GENERAL INFORMATION:  
; APPLICANT: ZUR MEGEDE, Jan  
; APPLICANT: BARNETT, Susan W.  
; APPLICANT: LIAN, Ying  
; APPLICANT: ENGELBRECHT, Susan  
; APPLICANT: VAN RENSBURG, Esmeralda J.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
; FILE REFERENCE: P018133.003 / 2302-18133  
; CURRENT APPLICATION NUMBER: US/10/190.435  
; NUMBER OF SEQ ID NOS: 319  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 59  
; LENGTH: 4713  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: tatRevNeiGagprotInaRmut C  
US-10-190-435-59

Query Match 99.7% Score 1474.8; DB 12; Length 4713;  
Best Local Similarity 99.9%; Pred. No. 2e-291;  
Matches 1476; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 TGGCGCGCCGCGCAGGATCTCGCGCGCGCGCAAGCTGGAGCGCTGGAGCGCATCCGCC 61  
Db 1262 TTGGCGCGCGCGCAGGATCTCTGGCGCGCGCAAGCTGGAGCGCTGGAGCGCATCCGCC 1321

QY 62 TGGCGCGCGCGCGCAGGATCTCTACATGATGAAGCACTGTGTGGCGCAGCGCCGAGC 121  
Db 1322 TGGCGCGCGCGCGCAGGATCTCTACATGATGAAGCACTGTGTGGCGCGCGCGGAGC 1381  
QY 122 TGGAGAAGTTCCGCTTGAACCCCGGCTGCTGGAGACAGCGAGAGGGCTGCAAGAGATCA 181  
Db 1382 TGGAGAAGTTCCGCTTGAACCCCGGCTGCTGGAGACAGCGAGAGGGCTGCAAGAGATCA 1441  
QY 182 TGGCGCAGCTGACCCCGCTGTCAGAGACCGGCGAGGAGGCTGAAGAGCCTGTTCACA 241  
Db 1442 TGGCGCAGCTGACCCCGCTGTCAGAGCGGCGAGGAGGCTGAAGAGCCTGTTCACA 1501  
QY 242 CCGTGCCCACTGCTACTGCTGTCAGAGAAATCGAGGTCCGCCACACCAAGAGAGGCC 301  
Db 1502 CCGTGCCCACTGCTACTGCTGTCAGAGAAATCGAGGTCCGCCACACCAAGAGAGGCC 1561  
QY 302 TGGACAAGATCAGGAGGAGCAGAACTGTCACAGAGATGCCAGAGATCCAGGAGCGCGG 361  
Db 1562 TGGACAAGATCAGGAGGAGCAGAACTGTCACAGAGATGCCAGAGATCCAGGAGCGCGG 1621  
QY 362 CCGACAAGGCAAGGTGAGCCAGAACTACCCCATCTGCAGAACTGCGAGGCGCCAGATGG 421  
Db 1622 CCGACAAGGCAAGGTGAGCCAGAACTACCCCATCTGCAGAACTGCGAGGCGCCAGATGG 1681  
QY 422 TGCACAGGCGCATCAGCCCGCCACCTGAACTGAGTGGTGAAGTGTGAGGAGAGAG 481  
Db 1682 TGCACAGGCGCATCAGCCCGCCACCTGAACTGAGTGGTGAAGTGTGAGGAGAGAG 1741  
QY 482 CTTGAGCGCCCGAGGTGATCGCCCATGTCACCGCTGAGCGAGGCGCCACCGCCAGG 541  
Db 1742 CTTGAGCGCCCGAGGTGATCGCCCATGTCACCGCTGAGCGAGGCGCCACCGCCAGG 1801  
QY 542 ACCTGAACACGATGTTGAACACCGTGGCGCGCCACAGCGCCCATGCGAGATGCTGAAG 601  
Db 1802 ACCTGAACACGATGTTGAACACCGTGGCGCGCCACAGCGCCCATGCGAGATGCTGAAG 1861  
QY 602 ACACCATCAACAGAGAGCGCGCGAGTGGAGCGGCTGCACCCCGTGCACGCGCGCCGCG 661  
Db 1862 ACACCATCAACAGAGAGCGCGCGAGTGGAGCGGCTGCACCCCGTGCACGCGCGCCGCG 1921  
QY 662 TGGCGCGCGCGCGAGTGGCGGAGCGCGCGGCGAGAGATCGCGCGCGCGCGCGCGCGCG 721  
Db 1922 TGGCGCGCGCGCGAGTGGCGGAGCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1981  
QY 722 TGCAGGAGCAGATCGCTGATGACCAAGCAACCCCGCCATCCCGCTGGCGCGCATCTACA 781  
Db 1982 TGCAGGAGCAGATCGCTGATGACCAAGCAACCCCGCCATCCCGCTGGCGCGCATCTACA 2041  
QY 782 AGCGTGGATCATCTGGGCTTGACAAAGATCGTGGGATGTACACCCCGCTGAGCATCC 841  
Db 2042 AGCGTGGATCATCTGGGCTTGACAAAGATCGTGGGATGTACACCCCGCTGAGCATCC 2101  
QY 842 TGGACATCAAGCAGGCG 901  
Db 2102 TGGACATCAAGCAGGCG 2161  
QY 902 CCGTGGCGCGCGCGAGCAGCACCCAGAGAGTGAAGAACTGAGTGAACCGACACCTGCTGG 961  
Db 2162 CCGTGGCGCGCGCGAGCAGCACCCAGAGAGTGAAGAACTGAGTGAACCGACACCTGCTGG 2221  
QY 962 TGCAGAACCCCG 1021  
Db 2222 TGCAGAACCCCG 2281  
QY 1022 TGCAGAGATGATGACCGCTGCCAGGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1081  
Db 2282 TGCAGAGATGATGACCGCTGCCAGGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2341  
QY 1082 TGGCGCGCGCGGATGAGCG 1141  
Db 2342 TGGCGCGCGCGGATGAGCG 2401  
QY 1142 GCGCGCGCGCGCATCTGTCAAAGTCTTCAACTGCGGCAAGAGGAGGCGCGCGCGCGCG 1201

```

|||||
Db 2402 GCGCGCGCGCATGCTCAAGTCTTCAACTGGGCAAGGAGGGCCACATCGCGCGCAACT 2461
QY 1202 GCGCGCGCGCGCGCAAGAGGCGTCTGGAATGCGCAAGGAGGGCCACACAGATGAAGG 1261
Db 2452 GCGCGCGCGCGCGCAAGAGGCGTCTGGAATGCGCAAGGAGGGCCACACAGATGAAGG 2521
QY 1262 ACCTGACCGAGGCGCGCAAGTTCCTGGGCAAGATCTGCGCCACGACCAAGGCGCGCG 1321
Db 2522 ACCTGACCGAGGCGCGCAAGTTCCTGGGCAAGATCTGCGCCACGACCAAGGCGCGCG 2581
QY 1322 CCGGCAACTTCCTGCAAGCGCGCGGAGCGCCACCGCGCGCGCGCGCGCGAGCTTCGCT 1381
Db 2582 CCGGCAACTTCCTGCAAGCGCGCGGAGCGCCACCGCGCGCGCGCGCGAGCTTCGCT 2641
QY 1382 TCGAGGAGACCGCGCGCGCGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1441
Db 2642 TCGAGGAGACCGCGCGCGCGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2701
QY 1442 TGAAGAGCGCTGTCGGCAAGCGCGCGCTGAGCGCAATA 1479
Db 2702 TGAAGAGCGCTGTCGGCAAGCGCGCGCTGAGCGCAATA 2739

```

Search completed: September 23, 2003, 02:00:04  
Job time : 375.76 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Computer Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 17:33:23 : Search time 95.5311 seconds  
(without alignments)  
5833.434 Million cell updates/sec

Title: US-09-475-704A-3

Perfect score: 1479

Sequence: 1 atggggcgccgcgcagcat.....acgaccccttgaccagtaa 1479

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 113956

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/p/odata/2/ina/5A-COMB.seq.\*  
2: /cgn2\_6/p/odata/2/ina/5B-COMB.seq.\*  
3: /cgn2\_6/p/odata/2/ina/6A-COMB.seq.\*  
4: /cgn2\_6/p/odata/2/ina/6B-COMB.seq.\*  
5: /cgn2\_6/p/odata/2/ina/PCFUS-COMB.seq.\*  
6: /cgn2\_6/p/odata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	947.6	64.1	4307	4	US-09-552-950-2
2	939.8	63.5	9772	4	US-09-552-950-5
3	738.2	49.9	1476	4	US-09-184-418C-100
4	738.2	49.9	8959	4	US-09-184-418C-11
5	729.8	49.3	1485	4	US-09-184-418C-73
6	729.8	49.3	9010	4	US-09-184-418C-8
7	710.8	48.1	1486	4	US-09-184-418C-38
8	710.8	48.1	8992	4	US-09-184-418C-4
9	701.2	47.4	1496	4	US-09-184-418C-82
10	701.2	47.4	8972	4	US-09-184-418C-9
11	659.8	44.6	1503	4	US-09-184-418C-29
12	659.8	44.6	8953	4	US-09-184-418C-3
13	658.8	44.5	1479	4	US-09-184-418C-12
14	658.8	44.5	8968	4	US-09-184-418C-1
15	656.6	44.4	4307	4	US-09-184-418C-48
16	656.6	44.4	8956	4	US-09-184-418C-5
17	652.6	44.1	1494	4	US-09-184-418C-47
18	650.2	44.0	4307	4	US-09-552-950-1
19	650.2	44.0	9719	4	US-09-700-304-1
20	647	43.7	3807	2	US-08-417-20A-78
21	645.4	43.6	7399	2	US-08-418-848A-9
22	645.4	43.6	9709	2	US-08-188-583-5
23	645.4	43.6	9709	3	US-08-388-353-1
24	645.4	43.6	9709	3	US-08-488-553B-1
25	645.4	43.6	9709	4	US-09-309-572-15
26	645.4	43.6	12479	4	US-09-318-138-13
27	645.4	43.6	12494	3	US-08-935-312-13

28	645.4	43.6	12494	3	US-08-848-760B-33	Sequence 33, Appl
29	645.4	43.6	15581	3	US-08-646-538-35	Sequence 35, Appl
30	645.4	43.6	15581	3	US-09-503-222-35	Sequence 35, Appl
31	642.2	43.4	1503	4	US-09-184-418C-64	Sequence 64, Appl
32	642.2	43.4	9050	4	US-09-184-418C-7	Sequence 7, Appl
33	641.4	43.4	8987	4	US-09-184-418C-2	Sequence 2, Appl
34	641	43.3	9737	2	US-08-944-449-7	Sequence 7, Appl
35	641	43.3	9737	2	US-09-353-362-7	Sequence 7, Appl
36	637	43.1	1491	4	US-09-184-418C-91	Sequence 91, Appl
37	637	43.1	9050	4	US-09-184-418C-10	Sequence 10, Appl
38	631.2	42.7	5362	3	US-08-463-20-5	Sequence 5, Appl
39	628	42.5	8932	3	US-09-124-900-1	Sequence 1, Appl
40	628	42.5	8933	3	US-08-463-210-4	Sequence 4, Appl
41	628	42.5	8933	3	US-09-620-958A-3	Sequence 3, Appl
42	628	42.5	8933	3	US-09-620-958A-4	Sequence 4, Appl
43	628	42.5	8933	3	US-09-620-958A-9	Sequence 9, Appl
44	624	42.2	9746	1	US-08-022-835-3	Sequence 3, Appl
45	624	42.2	9746	1	US-08-388-809-3	Sequence 3, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-552-950-2  
: Sequence 2, Application US/09552950  
: Patent No. 6541248  
: GENERAL INFORMATION:  
: APPLICANT: Oxford Biomedica (UK) Limited  
: TITLE OF INVENTION: Anti-Viral Vectors  
: FILE REFERENCE: 674524-2004  
: CURRENT APPLICATION NUMBER: US/09/552,950  
: CURRENT FILING DATE: 2000-04-20  
: NUMBER OF SEQ ID NOS: 22  
: SOFTWARE: PatentIn Ver. 2.1  
: SEQ ID NO 2  
: LENGTH: 4307  
: TYPE: DNA  
: ORGANISM: Artificial Sequence  
: FEATURE:  
: OTHER INFORMATION: Description of Artificial Sequence:gagpol-SYNgp - codon  
: OTHER INFORMATION: optimised gagpol sequence  
US-09-552-950-2

Query Match	64.1%	Score 347.6;	DB 4;	Length 4307;
Best Local Similarity	80.7%	Pred No. 9.2e-135;		
Matches 1135;	Conservative 0;	Mismatches 259;	Indels 12;	Gaps 2;
QY	1	ATGGCGCGCGCGCAGCATCTCTGGCGCGCGCAGCTGGAGCGCTGGAGCGGATCGCG	60	
DB	1	ATGGCGCGCGCGCAGCATCTCTGGCGCGCGCAGCTGGAGCGCTGGAGCGGATCGCG	60	
QY	61	CTGGCGCGCGCGCGCAGCATCTCTGGCGCGCGCAGCTGGAGCGCTGGAGCGGATCGCG	120	
DB	61	CTGGCGCGCGCGCGCAGCATCTCTGGCGCGCGCAGCTGGAGCGCTGGAGCGGATCGCG	120	
QY	121	CTGGAGAGTTTGGCGCTGAACCCCGCGCTCTGGAGCGCGCAGCGCTGGAGCGGATCGCG	180	
DB	121	CTGGAGAGTTTGGCGCTGAACCCCGCGCTCTGGAGCGCGCAGCGCTGGAGCGGATCGCG	180	
QY	181	ATCCGCCAGCTGCACCCCGCGCTCTGGAGCGCGCAGCGCTGGAGCGGATCGCG	240	
DB	181	ATCCGCCAGCTGCACCCCGCGCTCTGGAGCGCGCAGCGCTGGAGCGGATCGCG	240	
QY	241	ACCGTGGCGCGCGCTCTGGAGCGCGCAGCGCTGGAGCGGATCGCG	300	
DB	241	ACCGTGGCGCGCGCTCTGGAGCGCGCAGCGCTGGAGCGGATCGCG	300	
QY	301	CTGGAGAGATCGAGGAGGAGCAAGTCCAGCGAGAGATCCAGCGCGCGCGCGCG	360	
DB	301	CTGGAGAGATCGAGGAGGAGCAAGTCCAGCGAGAGATCCAGCGCGCGCGCGCG	360	
QY	361	GCC-----GACAGGGCAAGGTGAGCGAGCAACTATCCCATCTGTGCGAGACCTGCGGGC	414	

```

361 GACACCGGACACAGCAACAGGTCAGGCGAGAACTACCCCACTGGTCAGAAATCGAGGG 429
415 CAGATGTTGACACAGGATCAGCCCGGACCTGAACGCTGGTGAAGGTGATCGAG 474
421 CAGATGTTGACACAGGATCAGCCCGGACCTGAACGCTGGTGAAGGTGATCGAG 489
475 GAGAGGCTTCAGCCCGGAGGTGATCCCATGTTCAACGCTGGTGAAGGTGATCGAG 534
481 GAGAGGCTTCAGCCCGGAGGTGATCCCATGTTCAACGCTGGTGAAGGTGATCGAG 540
535 CCCAGAGACCTGAACAGATTTGAACACCTGGGCGGACACCGGCGGATCGAGATG 584
541 CCCAGATCTCAACACCATCTCAACACAGTGGGGGACACCGGCGGATCGAGATG 600
595 CTGAAGGACACCATCAACAGAGGCGCGGAGTGGGACCGCTGCACCGCTGCAGGCC 654
601 CTGAAGGACACCATCAACAGAGGCGCGGAGTGGGACCGCTGCACCGCTGCAGGCC 660
655 GGCCCATCGCCCGGCGGACATCGGCGAGTGGGCGGCGGACGATCGGCGGCGGAC 714
661 GGCCCATCGGCGGCGGACATCGGCGAGTGGGCGGCGGACGATCGGCGGCGGAC 720
715 AGCACCTTCAGGAGGACATCGCTGATGATGACAGCAACCGGCGGCGGAG 774
721 AGTACCTTCAGGAGGACATCGCTGATGATGACAGCAACCGGCGGCGGAG 780
775 ATCTACAGCGGTGGATCATCTCGGCGTGAACAGATCGTGGGATGTACAGCGCGGTG 834
781 ATCTACAGCGGTGGATCATCTCGGCGTGAACAGATCGTGGGATGTACAGCGCGGTG 840
835 AGCATCTGGACATCAAGACGGGCGGCAAGAGGCGCTTCGCGACTACGTTGAGCGTTC 894
841 AGCATCTGGACATCAAGACGGGCGGCAAGAGGCGCTTCGCGACTACGTTGAGCGTTC 900
895 TTCAAGACCGTTCGCGGCGGAGGACGAGGACGAGGAGTGAAGACTGATGACGAGCAC 954
901 TACAAACCGTTCGCGGCGGAGGACGAGGAGTGAAGACTGATGACGAGCAC 960
955 TGCTGTGTGAGAGGACGAGGACGAGGAGTGAAGACTGATGACGAGCAC 1014
961 TGCTGTGTGAGAGGACGAGGACGAGGAGTGAAGACTGATGACGAGCAC 1020
1015 GCCAGCTTGAAGGACGAGGACGAGGAGTGAAGACTGATGACGAGCAC 1074
1021 GCTACCTTGAAGGACGAGGACGAGGAGTGAAGACTGATGACGAGCAC 1080
1075 CGCTGTGTGAGGAGGACGAGGACGAGGAGTGAAGACTGATGACGAGCAC 1128
1081 CGCTGTGTGAGGAGGACGAGGAGTGAAGACTGATGACGAGCAC 1140
1129 AGCAACTTGAAGGACGAGGACGAGGAGTGAAGACTGATGACGAGCAC 1184
1141 GCGAACTTGAAGGACGAGGACGAGGAGTGAAGACTGATGACGAGCAC 1200
1189 ATGCGCGCACTGAGGAGGACGAGGAGTGAAGACTGATGACGAGCAC 1248
1201 ACAGCGCGCACTGAGGAGGACGAGGAGTGAAGACTGATGACGAGCAC 1260
1249 CACAGATGAAGGACTGAGGAGGACGAGGAGTGAAGACTGATGACGAGCAC 1308
1261 CACAGATGAAGGACTGAGGAGGACGAGGAGTGAAGACTGATGACGAGCAC 1320
1309 CACAGGCGCGCGGCACTTCTGAGAGCGCGCGGAGGAGTGAAGACTGATGACGAGCAC 1368
1321 TACAGGAGGAGGAGGAGGAGTGAAGACTGATGACGAGCAC 1380
1369 GAGAGCTTCCGCTTCGAGGAGGAGGAGGAGTGAAGACTGATGACGAGCAC 1394
1381 GCGAGCTTCAGGCTTCGAGGAGGAGGAGGAGTGAAGACTGATGACGAGCAC 1406
```

RESULT 2

```

US-09-552-950-5
; Sequence 5, Application US/09552950
; Patent No. 6541248
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Limited
; TITLE OF INVENTION: Anti-Viral Vectors
; FILE REFERENCE: 674524-2004
; CURRENT APPLICATION NUMBER: US/09/552,950
; NUMBER FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pSYNG?
US-09-552-950-5
```

```

Query Match      63.5%; Score 939.8; DB 4; Length 9772;
Best Local Similarity 78.6%; Pred. No. 1.3e-133;
Matches 1182; Conservative 0; Mismatches 297; Indels 24; Gaps 4;

QY      1 ATGGGCGCGCGCGCGCGAGCATCTCTGGCGCGCGGCGGCGGAGCTGGAGCGCTGGAGCGCATCCGC 60
DB      1108 ATGGGCGCGCGCGCGCGAGCGTGTCTGGCGCGCGGAGCTGGAGCGCTGGAGCGCATCCGC 1167

QY      61 CTGCGCGCGCGCGCGGCAAGAAAGTCTACATGATGAAGCACCTGTGTGTGGCGCGGCGCGGAG 120
DB      1168 CTGCGCGCGCGCGCGGCAAAAGAAAGTACAAAGCTGAAGCACATGTGTGTGGCGCGGCGGAA 1227

QY      121 CTGGAAGAGTTCGCGCTGAACCGCGCGCTGTGTGGAGACCGAGGAGGTGCAAGCAGATC 180
DB      1228 CTGGAGCGCTTCGCGCTGAACCGCGCGCTGTGTGGAGACCGAGGAGGTGCGCCCGCAGATC 1287

QY      181 ATCGCGCAGCTGCACCGCGCGCGCTGTGTGGAGACCGCGGCGGAGGAGCTGAAGAGCGTTCAC 240
DB      1288 CTGGCGCAACTGCAGCGCGCGCGCTGTGTGGAGACCGCGGAGGAGCTGCGCAGCGCTGTACAA 1347

QY      241 ACCTGTGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
DB      1349 ACCTGTGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1407

QY      301 CTGGACAGATTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 354
DB      1408 CTGGATAAAATCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1467

QY      355 GAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 414
DB      1468 GAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1527

QY      415 GAGATGCTGCACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 474
DB      1528 CAGATGCTGCACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1587

QY      475 GAGAGCGCTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 534
DB      1588 GAGAGCGCTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1647

QY      535 CCGCAGGAGCTGAACAGCATGTGTGAACCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
DB      1648 CCGCAGATGTGAACAGCATGTGTGAACCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1707

QY      595 CTGAAGGAGCAGCATCAACGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 654
DB      1708 CTGAAGGAGCAGCATCAACGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1767

QY      655 GGCGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 714
DB      1768 GGCGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1827

QY      715 AGCACCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 774
DB      1827 AGCACCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1886
```

```
Db 1828 AGTACGCTTCAGGAACAGATCGCTGGATGACCAACACACCCACCCATCCCGGTGGAGAA 1887
QY 775 ATCTACAGCGTGGATCATCTCTGGGCTTGAAACAAATCGTGGATGTACAGCCCGCTG 834
Db 1888 ATCTACAAAGCGTGGATCATCTCTGGGCTTGAAACAAATCGTGGATGTACAGCCCGT 1947
QY 835 AGCATCTGGACATCAAGCAGGCGCTCAAGAGAGCGCTTCGGGACATACGTGGAGCGCTTC 894
Db 1948 AGCATCTGGACATCGGCAAGCGCGCGAGAGAGCGCTTCGGGACATACGTGGAGCGCTTC 2207
QY 895 TTCAAGACGCTCGGCGCGAGAGAGAGCGCGAGAGAGCGCTTCGGGACATACGTGGAGCG 954
Db 2208 TACAAAACGCTCGGCGCGAGAGAGCGCTTCGGGACATACGTGGAGAGCGCTTCGGGAC 2057
QY 955 CTGCTGCTGAGAGAGAGCGCGAGAGAGCGCTTCGGGACATACGTGGAGAGCGCTTCGG 1014
Db 2068 CTGCTGCTGAGAGAGAGCGCGAGAGAGCGCTTCGGGACATACGTGGAGAGCGCTTCGG 2127
QY 1015 GCCAGCTGAGAGAGATGATGAGCGCGCTGCGAGAGAGCGCGCGCGCGCGCGCGCGCG 1074
Db 2128 GCTACCTAGAGGAATGATGACCGCGCTGTCAGAGAGAGCGCGCGCGCGCGCGCGCG 2187
QY 1075 CGGCTGCTGCGCGAGAGAGCGCGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1124
Db 2198 CGGCTGCTGAGGCGCATGAGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2247
QY 1129 AGCAAGCTCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1188
Db 2248 GCGAACTTCGGAAACAGCGCAAGATGCTCAAGTGTCTTCAAGTGTCTTCAAGTGTCTT 2307
QY 1189 ATGCGCGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1248
Db 2308 ACAGCGCGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2367
QY 1249 CACCAATGAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1308
Db 2368 CACCAATGAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2427
QY 1309 CACAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1368
Db 2428 CACAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2487
QY 1369 GAGAGCTTCGCTTCAGGA-----CAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1416
Db 2488 GAGAGCTTCAGGTTTGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2547
QY 1417 AAGGACCGGAGAGCGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1476
Db 2548 AAGGAACTGATCTTATGCTTCCTTCAGATCACTCTTTGGCAGCGAGCGCGCGCGCGCG 2607
QY 1477 TAA 1479
Db 2608 TAA 2610
```

## RESULT 3

```
US-09-184-418C-100
; Sequence 100, Application US/09184418C
; Patent No. 649210
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 100
; LENGTH: 1476
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
```

```
FEATURE:
OTHER INFORMATION: isolate=94IN476.104; gene=gag
US-09-184-418C-100
Query Match 49.9%; Score 738.2; DB 4; Length 1476;
Best Local Similarity 69.2%; Pred. No. 2.7e-103;
Matches 1023; Conservative 0; Mismatches 453; Indels 3; Gaps 1;
QY 1 ATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db 1 ATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 61 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 TTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 CTGGAGAGTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 CTGGAAGATTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 181 ATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 ATAAACACGCTACATCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 241 ACCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 241 ACAGTAGCAACTCTCTATTGTGTACATGCGAGGATAGAGTACGAGACACCAAGAGCGC 300
QY 301 CTGGACAGATCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 301 TTAGACAAGATAGAGGAAGACAAACAAAGTCAAGCAAAATACAGCAGGCAAGAGAG 360
QY 361 GCGGACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Db 361 GCTGAC---GGAAAGGTCAGTCAAAATATCTATAGTGCAAAATCTCCAGGGCAATG 417
QY 421 GTGCACGAGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 418 GTACACGAGCGCGCATCAGCTAGACTTTGATGCGTGGTAAAGTAATAGAGGAGAG 477
QY 481 GCTTTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 478 GCTTTAGCGCGCGCGCATATACCGCATGTCACAGCATATCAGAGAGAGCGCGCGCGCG 537
QY 541 GACCTGACACCATCTTGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 538 GATTTAAACACCATCTTAATACAGTGGGGGAGCATCAGAGCGCGCATGCAATGTTTAAA 597
QY 601 GACACCATCAACGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 598 GATACCATCAACGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657
QY 661 ATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Db 658 AATCCACGAGCGCGCATGAGAACCAAGGGAAGTGATATAGCAGGAACTACTAGTACC 717
QY 721 CTGAGGAGCGCATCGCTGGATGACCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
Db 728 CTTCAGGAACAATACCATGGATGACAGTACCCACCGCTATTCAGTGGGAGCATCTAT 777
QY 761 AAGCGGTGATCATCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
Db 778 AAAAGATGATAATCTGGGGTTAAATATAAATAGTAAGAAATGATATAGCGCTGTCAGCAT 837
QY 841 CTGGACATCAACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db 838 TTGGACATAAGACAAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897
QY 901 ACCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
Db 898 ACTTAAAGCTGGAACAAGCTACACAAGAAAGTAAAGGTTCGATGACAGACCGCTGTG 957
QY 961 GTGCAGAGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
```

```

Db      958  GTCCAAATGCAAAACCCAGATGTAAGACCATTTAAGACCA:TAGGACCCAGGGCTTCA 1:17
Qy      1021  CTGGAGAGATGATGACGCTCTGCCAGGGGTGGGCGGCCCCACCAACAGGCGCGGCTG 1:80
Db      1018  TTAGAAGAAATGTTGACAGCATGTTCAAGGAGTGGGAGGACTAGCCCAAGCAAGAGT 1:77
Qy      1081  CTGCGCCGAGGATGATGACGCTGACGCAACCCAGGCGTGAIGATGAGAGAGCAACTTCA 1:40
Db      1078  TTGGCTGAGCAATGAGCCCATACATAGTAAACATAAIGATGACAGAGAGCCATTTAA 1:17
Qy      1141  GGCCCCGGCGCATGCTCAAGTGTCTCAACTGCGGCAAGGAGGCGCACATCGGCCGCA 1:20
Db      1138  GCCCCTAAAGAAATGTTAAAGTCTCAACTGTGGCAAGGAGGCAATAGCCAGAAAT 1:97
Qy      1201  TGCCGCGCCCCCAGCAAGAGGGTGTCTGGAAGTGGCGGCAAGGAGGCGCCACCATGA 1:20
Db      1198  TGCAGGCGCCCTAGAAAAGAGGTGTGGAAATGTGGGCAAGGAGGACCAATGAAA 1:27
Qy      1261  GACTGCAACGAGCGCCAGCCCAACTTCTGGGCAAGATCTGGCCAGCCCAAGGGGCGC 1:20
Db      1258  GACTGTACTGAGAGCGAGCTAATTTTATAGGGAATTTGGCCTTCCCAAGGGGAGG 1:17
Qy      1321  CCCGGCAACTTCTGCAAGAGCGCCCGGAGCCCAACCGCCCCCGCGGAGAGCTTCGC 1:30
Db      1318  CCAGGGAATTCCTTCAAAACAGCCAGAGCCCAACAGCCCAACAGGAGAGCTTCAG 1:37
Qy      1381  TTCAGAGAGACCCCGGCGCAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1:40
Db      1378  TTCAGGAGAGCAACCCCGCTCCGAGGAGAGTTCGAAAGCAGGAGGAGGAGGAGGAG 1:37
Qy      1441  CTGAAGAGCGCTTTCGGCAAGCAACCGCTGAGCCAGTAA 1:79
Db      1438  CTCAAATCACTCTTTGGCAGGCGCCCTTGCTCAATAA 1:46

```

## RESULT 4

```

US-09-184-418C-11
; Sequence 11, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184.418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 11
; LENGTH: 8959
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURES:
; OTHER INFORMATION: isolate=941N476.104; 138.1613;"aaq";
; OTHER INFORMATION: 1418.4428;"pol"; 4361.4939;"vif"; 4879.5169;"vpr";
; OTHER INFORMATION: 5150.7782;"tat"; 5289.7939;"rev"; 5378.5638;"vpu";
; OTHER INFORMATION: 5556.8129;"env"; 8131.8754;"ref"
US-09-184-418C-11

```

```

Query Match      49.9%   Score 718.2; DB 4; Length 8959;
Best Local Similarity 69.2%   Pred. No. 2.6e-103;
Matches 1023; Conservative 0; Mismatches 453; Indels 3; Gaps 3;

```

```

Qy      1  ATGGCGCGCGCGCCAGCATCTCGGCGGCGGCAAGCTGGAGCGCTGGAGCGGATCCG 60
Db      138  ATGGGTCGAGAGCGGTCAATATTAGAGGGGGAAATTAGAATAGGGGAAAAATCGG 197
Qy      61  CTGCGCGCGCGCGGCAAGAGTGTACATGATGAGCACCTGGTGTGGCGGCGAGCGGAG 120
Db      198  TTAGGCGCGCGGGGAGAAACATATTATATGATTAACACTTAGTATGGCAAGCGGGAG 257

```

```

Qy      121  CTGGAGAGTTGCCCTGAACCCCGGCTCTGGAGACCAAGGAGGCTGCAAGCAGATC 180
Db      258  CTGGAAGATTTGGCGTTAAACCTGGCTTTTAGAGACCTCAGACGGATGTAAACAATA 317
Qy      181  ATCCGCGAGCTGACCCCGCTGACAGCGGAGGAGGCTGAAGASCTGTGTTCAAC 240
Db      319  ATAAACACAGCTACATCAGCTCTTTAGACAGGAGACAGAGAACTTAGTTCATTATCA 377
Qy      241  ACGTGGCCACCTGTACTGGTGCACGAGAGATGAGGTCCGCGACACCAAGGAGGCG 300
Db      378  AAGTAGCAACTCTCTATTGTGTACATGCGAGGATAGAGTAGGACACCAAGGAGGCC 437
Qy      301  CTGGCAAGATCGAGGAGGAGCAAAAGTSCCAGCAGAAATCCAGAGCGCCGAGGCC 360
Db      438  TTAGCAAGATAGAGGAGAGCAAAAGTTCAGCAAAATAACAGCAGGCAAAAGAG 497
Qy      361  GCGGCAAGGAGGAGTGAAGCCAGAACTACCCATCGTCAGAACTCTGAGGCGCGATG 420
Db      498  GCTGAC---GGAAAGGTCAGTCAAAATATCCCTATAGTCAAAATCTCCAAAGGCAATG 554
Qy      421  GTGCACAGGCGCATCAGCCCGCCACCGCTGAAGCGCTGGTGAAGGTGATCGAGGAGAG 480
Db      555  GTACACAGCCCTATCACCTAGAACTTTGAATTCGCTGGTAAAGTAAATAGAGGAGAG 614
Qy      481  GCCTTCAGCCCGGAGGTGATCCCGCATSTTCAACCGCCCTGAGCGAGGCGCCACCC 540
Db      615  GCTTTTAGCCAGAGGTAAATACCATGTTACAGAGCATTATCAGAGGAGGCCACCCCT 674
Qy      541  GACCTGAACACAGTGTGAACACCGTGGGGGCGCCACAGCGCGCCATCCAGATGCTGA 600
Db      675  GATTTAACACCATGTTAAATACAGTGGGGGAGACATCAAGCAGCGCATGCAAAATGT 734
Qy      601  GACACCATCAACGAGGAGCGCGCGAGTGGGACCGCTGCACCGCTGCACCGCGCGCC 660
Db      735  GATACCATCAACGAGGAGGCTGGGAATGGGATAGATTACATCCATACATCGAGGCGCT 794
Qy      661  ATCGCCCGCGGCGAGATGCGCGAGCGCGCGGAGCATCGCGGAGCATCGCGGAGCAT 720
Db      795  AATCCACAGCGCGAGATGAGAGAACCAAGGGAAGTGAATAGCAGGAACTACTAGTAG 854
Qy      721  CTGCAGGAGCAGATCGCTGGATGACCAACCAACCCCGCATCCCGTGGGAGCATCTAC 780
Db      855  CTTGAGGAAACAAATAGCATGGATGATGACAGGTAACCCACCTATTCCAGTGGAGACA 914
Qy      761  AAGCGGTGGATCATCTCGGCGCTGAACAGATCGTGGGATGTACAGCCCGCTGAGCAT 840
Db      915  AAAAGATGGATAATTCGGGTTAAATAAATAGTAAGAATGTATAGCCCTGTCAGCAAT 974
Qy      841  CTGGACATCAAGCAGGCGCCCAAGGAGCGCTCCGCGAGTACGTGGAGCGCTTCTTCA 900
Db      975  ITGGACATAAGACAGGCGCCCAAGGAAACCGCTTTAGAGACTATGTAGACCGGTTCTT 1034
Qy      901  ACCCTGCGCGCGAGCAGACACCGAGAGGTGAAGAAGTGGATGACCGACACCGCTGCT 960
Db      1035  ACTTTAGAGCTGACACAGCTACACAGAGTAAAGGTTGAAGACACACACTTGTG 1094
Qy      962  GTGAGAACGCCAACCCCGACTGCAAGACCATCTCGCGCTCTCGGCGCGCGCGCAGC 1020
Db      1095  GTCCAAAATCAACCCAGATCTGAAGACCATTTAAGAGCATTAGGACCGAGGCGCTTCA 1154
Qy      1022  CTGGAGAGATGATGACCGCTGCCAGGGGTGGCGGCGCCAGCCAGCGCCGCGGCTG 1080
Db      1155  TTAGAAGAAATGGTGACAGCATGTCAAGGAGTGGAGGAGCTAGCCCAAGGAGAGTG 1214
Qy      1081  CTGGCGAGCGGATGAGCCAGGCGCCACACCGCTGATGATGAGAGAGGAGCACTTCA 1140
Db      1215  TTGGCTGAGGCAATGAGCCCAATCACATAGTAACATAATGATGAGAGAGGCAATTTAA 1274
Qy      1141  GGCCCCCGCGCGCTCGTCAAGTGTCTCAACTGGGGAAGGAGGCGCCACATCGGCCCA 1200
Db      1275  GGCCCTAAAGAAATGTTAAATGCTTCAACTGTGGCAGGAGGCGCACATAGCCAGAA 1334
Qy      1201  TGCCGCGCGCGCGCGCAAGAGGCGCTCTGGAAGTGGCGGAAGGAGGCGCCAGCATGA 1260

```



1335 TGCAGGGCCCTAGAAAAAGAGGCTCTTGGAAATGTGGCCAAAGGACACCAATGAAA 1394  
 1261 GACTGCAGGAGCGCCAGCGCAACTTCCCTGGGCAAGATCTGGCCACAGCCACAGGGCCGC 1320  
 1395 GACTGTACTGAGAGGAGGAGCTTAATTTTGGGCAAAATTTGGCTTCCCAAGGGAGG 1454  
 1321 CCGGGCAACTTCTGAGAGCGCGGAGCGCCAGCCACCGCCCGCGGAGAGGTTCCGC 1380  
 1455 CAGGGAAATTCCTTCAAAACAGCCAGAGCCCAACAGCCCCACAGAGAGAGCTCAGG 1514  
 1381 TTGAGGAGACCAACCCCGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440  
 1515 TTCAAGGAGACAAACCCCGCTCGGAAGCAGGAGTGGAAAGACAGGGAACCCCTTAACCTTC 1574  
 1441 CTGAAGAGGCTGTTCGGCAACGACCCCTGAGCCAGTAA 1479  
 1575 CTCAAATCACTCTTTGGCAGCGACCCCTTGTCTCAATAA 1613

## RESULT 5

US-09-184-418C-73  
 : Sequence 73, Application US/09184418C  
 : Patent No. 6492113  
 : GENERAL INFORMATION:  
 : APPLICANT: Hahn, Beatrice  
 : APPLICANT: Gao, Feng  
 : APPLICANT: Shaw, George  
 : TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN  
 : TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1  
 : FILE REFERENCE: D6287  
 : CURRENT APPLICATION NUMBER: US/09/184,418C  
 : CURRENT FILING DATE: 1999-11-02  
 : NUMBER OF SEQ ID NOS: 112  
 : SEQ ID NO 73  
 : LENGTH: 1485  
 : TYPE: DNA  
 : ORGANISM: Human immunodeficiency virus type 1  
 : FEATURE:  
 : OTHER INFORMATION: isolate=962M651.8; gene=gag  
 US-09-184-418C-73

Query Match 49.3%; Score 729.8; Db 4; Length 1485;  
 Best Local Similarity 69.2%; Pred. No. 4,9e-102;  
 Matches 1027; Conservative 0; Mismatches 452; Indels 6; Gaps 2;

1 ATGGGGCGCGCGCCAGCATCTGCGCGGCGCAAGCTGGAGCCCTGGAGCGCATCCGC 60  
 1 ATGGGTGCGAGAGCGTCAATATTAGAGGGGGAATTTAGTAAATGGGAAAAATTAGG 60  
 61 CTGCGCGCGCGCGCGCAAGAGTGTACATGATGAAGCAGCTGGTGTGGCCAGCGGGAG 120  
 51 CTAAAGCCAGGGGAAAGAAACGCTATATGATAAACACCTAGTATGGCAAGCAGGGAG 120  
 121 CTGGAGAAAGTTCGCTGAAACCCCGCTGCTGGAGACAGGAGGGCTGCAAGCAGATC 180  
 121 CTGGAAGAATTTCGCTTAACCTTGCCCTTTAGAAACATCAGAGGCGCTTAACAAATA 180  
 181 ATCGGCGAGCTGCACCCCGCTGCGAGACGCGGAGGAGGAGTGAAGCGCTGTTCAC 240  
 181 ATGAACACGCTACACCGAGCTTTCAAGAGGAATGGAGGAGAACTTAGATCAATTACAA 240  
 241 ACGGTGGCCACCGTGTACTGGGTGACGAGAGATCGAGGTCCGGACACCAAGAGAGGTC 300  
 241 ACAGTAGCAACTCTCTATGTGTACATGAAGGGTAGAGGTACGAGACACCAAGNACCC 300  
 301 CTGGACAAAGATCGAGGAGGAGAGCAACAAGTGCAGCAAGAGATCTAGC---AGGCCGAG 357  
 301 TTAGACAGGATAGAGGAAGAACAAACAAATTCAGCAAAATACAGCAAAATACAGAAAAACAG 360  
 358 GCGCGCGCAAGGGCAGGTGAGCCAGAACTACCCCATCTGCAACACTTCAGAGGCCAG 417  
 361 CAAGCGGCTGACGGAAAGGTGCTCAAAATTTATCTATAGTGCAGAACTCTCAAGGGTAA 420

418 ATGGTGCACAGGCCATCAGCCCGCCGACCCCTGAACCCCTGGTGAAGCTGATCGAGGAG 477  
 421 ATGTGTACACAGAAACTATCACTAGAACTTTGAATGCTATGGGTAAAGTAAATAGAAAGA 480  
 478 AAGGCGCTTCAGCCCGGAGGTGATCCCATGTTACCGCCCTGAGCGAGGGCGCCACCCGC 537  
 481 AAAGCTTTAGCCACAGAGGTAAATACCATGTTTACAGCAATTATCAGAGAGGACACCCCA 540  
 538 CAGGACCTGAACAGATGTTGAACACCGTGGCGGCCACACAGGCCGCTATCGACAGATCG 597  
 541 CAAGATTTAAACACCACTGTTAATATACATGGGGGAGCATCAAGCAGCATGCAAAATGTA 600  
 598 AAGGACACCATCAAGGAGGAGCGCGAGTGGACCGGTCACCGGTCACCGCTGCACGCCGCG 657  
 601 AAGATCTATCAATGAGAGGCTGCAAGTGGATAGATTACATCCAGTGCATGCGAGGG 660  
 656 CCGATCGCCCGCGGAGATCGGCGAGCGCGCGGAGCGACATCGCGCGGCGCACCCAGC 717  
 661 CCTATTGCACAGGCCAAATGAGAGAACCAAGGGAAAGTGATATAGCAGGAATCTAGT 720  
 718 ACCCTGAGGAGCAGATCGCCTGATGACACAGCAACCCGCCCATCTCCGCTGGCGGACATC 777  
 721 ACCCTCCAAAGAACAGATAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 780  
 778 TACAAGCGGTGGATCATCTCGGCTGTAACAAGATCGTGGGATGTACAGCCCGGAGC 837  
 781 TATAAAGATGATTAATCTGGGTAAATAAATAGTAAATAGTAAATAGTAAATAGTAAATAG 840  
 838 ATCTTGGACATCAAGAGGCGCCCAAGAGCGCTTCGCGGAGTACGTGGAGCGGTCTTC 897  
 841 ATTTTGGACATAAACAAGGCGCAAGAACCTTTAGAGACTATAGAGCGGTCTTC 900  
 898 AAGACCCCTGCGCGGAGCAGCAGCAGCCGAGGTGAAGAACTGGATGACCGACCCG 957  
 901 AAAACITTAGAGCTGAACAGGCTACACAAAGATAAATAATGGATGACAGACACCTTG 960  
 958 CTGTTGCAAGACGCAACCCGAGCTGCAAGACCATCTCTGCGCCCTCTCGGCGCCCGGCGCC 1017  
 961 TTGGTCCAAATGCAACCCAGATGCAAGACCATTTTAAAGCAATTAGGACCGAGGCT 1020  
 1018 AGCTTGAGGAGATGATGACCGCTGCGAGGGGTGGCGGCCCGCCAGCCAGCGCCGCG 1077  
 1021 ACATTAGAAATATGATGACAGCATGTCAAGGAGTGGGAGGACCTAGCCCAAGAGCAAG 1080  
 1078 GTGTCGCGGAGGCGATGAGCGAGCGCAACA---CCAGCTGATGATGCAAGAGCAAC 1134  
 1081 GTGTTGGCTGAGGCAATGAGCCAAACAAATAGTGTAAACATCTGATGCGAAAGCAAT 1140  
 1135 TTCAAGGGCGCCCGCGCATGCTCAAGTGCCTCAACTCGCGCAAGAGGGGCGACATCGCC 1194  
 1141 TTAAAGGAAATAAAGAAATGTTAAATGTTTAACTGTGTGTAAGGAAGCGCACATAGCC 1200  
 1195 GCAACTGCGCGCGCCCGCGCAAGAGGCTGTGAGTGGCGCAAGAGGAGGCGCCAG 1254  
 1201 AGAAATTCGAGGCGCCCTAGAAAAAGGGCTGTGGAATGTGGAAGAGGAGGAGACAA 1260  
 1255 ATGAAGGACTGCACGAGCGCGCGCAACTCTCTGGCAAGATCTGGCCAGCCACAAG 1314  
 1261 ATGAAGACTGCTACTGAGAGCGAGCTAATTTTATAGGAAATTTGGCGCTTCCACAG 1320  
 1315 GGCGCGCCCGGCAACTTCTCTGAGAGCGCGCCGAGCGCCAGCGCCCGCCCGCGGAGC 1374  
 1321 GGAAGGCGAGGAAATTTCTTCAACAGCGCCAGAGCCACAGCGCCCGCCAGCAGAGC 1380  
 1375 TTGCGCTTGGAGGAGACACCCCGCGCGAGAGCAGGAGAGCAAGAGCCGCGGAGCCCTG 1434  
 1381 TTCAGTTCGAGGAGACAAACCCCGCTCCGAAGCAGGAGTCAAGAGAGCAGGAGGAGCTTA 1440  
 1435 ACCAGCTGAGAGCGCTGTTGCGCAAGCAGCCCTGAGCCAGTAA 1479  
 1441 ACTTCCCTCAATCACTCTTTGGCAGCGACCCCTGCTCTCAATAA 1485

## RESULT 6

US-09-184-418C-8  
 : Sequence 8, Application US/09184418C  
 : Patent No. 6492110  
 : GENERAL INFORMATION:  
 : APPLICANT: Hahn, Beatrice  
 : APPLICANT: Gao, Feng  
 : APPLICANT: Shaw, George  
 : TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN  
 : FILE REFERENCE: D6287  
 : CURRENT APPLICATION NUMBER: US/09/184,418C  
 : CURRENT FILING DATE: 1999-11-02  
 : NUMBER OF SEQ ID NOS: 112  
 : SEQ ID NO 8  
 : LENGTH: 9010  
 : TYPE: DNA  
 : ORGANISM: Human immunodeficiency virus type 1  
 : FEATURE:  
 : OTHER INFORMATION: isolate=962M651; 137..1621:"gag"; 1426..4425:"pol";  
 : OTHER INFORMATION: 4370..4548:"vif"; 4888..5178:"vpr";  
 : OTHER INFORMATION: 5159..5373-7734..7824:"tat"; 5298..5373-7734..7581:"rev";  
 : OTHER INFORMATION: 5387..5647:"vpu"; 5565..8171:"env"; 8173..8793:"nef";  
 : US-09-184-418C-8

Query Match 49.3%; Score 729.8; DB 4; Length 9010;  
 Best Local Similarity 69.2%; Pred. No. 4.7e-102;  
 Matches 1027; Conservative 0; Mismatches 452; Indels 5; Gaps 2;

QY	1	ATGGGGCGCCGCGCCAGCATCTTGGGGGGGGAAGCTGGAGCCCTGGGGAGCGCATCGGC	50
DB	137	ATGGGTGCGAGAGCGTCAATTAAAGAGGGGAAATTAGATAAATGGGAAATAATAGG	196
QY	61	CTGGCGCCCGCGCCAAAGAGTGCTACATGATGAAGCACCTGTGTGGCGACCGCCGAG	120
DB	197	CTAAGCCAGGGGAAAGAACCTATATGATAAACACTATATATGGCGAAGCGGAG	256
QY	121	CTGGAGAAGTTCCCTTGAAACCGGCTCTCTGGAGACCGAGGGCTGCAAGCAGATC	180
DB	257	CTGGAAGATTTGGCTTAAACCTGCGCTTTAGAAACATCAGAAGGCTTAAACAAATA	316
QY	181	ATCGGCAGCTGACCCCGCCCTGCAAGCCGCGAGGAGAGTGAAGAGCCCTSTTCAGC	240
DB	317	ATGAACAGCTACAACTGCTCTGACAGAGGAGGAGGAGGAGTATGATGATATACAA	376
QY	241	ACGCTGCCACGCTGTACTTCGTGGCAGAGAAGATCGAGTCCGCGACACCAAGAGGCG	300
DB	377	ACAGTACCACTCTCTATCTGTATGATGAAGGGTAGNGTACGACACCAAGAGAGCT	436
QY	301	CTGGACAGATTCAGGAGGAGGAGAGACAGATGCGCAGAGAGTCTAGC---AGGCGGAG	357
DB	437	TTAGACAGGATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	496
QY	358	GCCGCGCAGACAGGCGAAGTGAAGCAGAGTACCCATCGTGGCAGAGGAGGAGGAGG	417
DB	497	CAGCGGCTGACGGAAGGTCAGTCAAAATATTCCTATAGTGCAGATCTCCAGGCGCA	556
QY	418	ATGGTGACAGGCCCATCATCCCGCCGACCCCTGAAGCGCTGGGTGAAGTGATGAGGAG	477
DB	557	ATGGTACACCAAGAACTATCAACCTGAAGACTTTGAATGCAAGGCTGAAGATTAAGAGAA	616
QY	478	AAGGCTTACGCGCGAGGTGATCCCATGTTTACCGGCTGAGCGAGGCGGCGAGCGCG	537
DB	617	AAAGCTTTAGCCCGAGAGTAAACCCATGTTTACAGCATTTACAGAGGAGGAGGAGGAG	676
QY	538	CAGAGCTGAACAGATCTTGAACAGCTGGCGGCGACAGGAGGAGGAGGAGGAGGAGG	597
DB	677	CAGATTTAAACACCATGTTAAATACAGTGGGGGACATCAAGCGCCATGCAAAATGTTA	736
QY	598	AAGGACACCATCAAGGAGGCGCGAGTGGGAGCGGCTGACCGGCTGACCGCGGCGGCG	657
DB	737	AAAGATACTATCATGAGGAGGCTGCAGATGGATATGATATCATCCAGTGCATGAGGAG	796

QY	658	CCCATGGCCCGCGCCAGATGCGCAGGCGCCGCGAGCGACATCCCGCGGACACACGAG	717
DB	797	CTCTATGCGACAGGCGCAATAGAGAGAACCAAGCGGAGTATATACAGGAATCTACT	856
QY	718	ACCCTGCGAGGAGCAGATCGCCCGATGATGACCAACACACCCCGCATCCCGTGGGAGACAT	777
DB	857	ACCTCCAGAGAACAGATAGCATGATGATGATGATGATGATGATGATGATGATGATGATG	916
QY	778	TACAGCGGTGATCATCTGCGGCTGACACAGATCGTGGGATGATGATGATGATGATGATG	837
DB	917	TATAAAGATGATTAATCTGGGGTTAAATAAATAGTAAGAAATGATAGCCCTCTCAGC	976
QY	838	ATCCTGGACATCAAGCAGGCGCCCAAGAGCGCTTCCCGCATGATGATGATGATGATGATG	897
DB	977	ATTITGGACATAAACAAGGCGCAAGGAGCGCTTTAGAGACTATGATGATGATGATGATG	1036
QY	898	AAGACCTGCGCGCGCGAGCAGAGCAGCAGGAGGTGAAGAACTGGATGATGATGATGATG	957
DB	1037	AAAACCTTTAAGAGCTGAACAGGCTACACAGAGTAAATAAATTTGGATGATGATGATG	1096
QY	958	CTGGTGCAGAGCGCAACCCCGACTGCAAGACCATCTCGCGGCTCTCGGCGCGCGCGCG	1017
DB	1097	TTGGTCCAAATGCAACCCAGATTCAGAGACCATTTTAAAGACATTAGGACCGGCTCT	1156
QY	1018	AGCCTGGAGGAGATGATGACCGCTGCCAGGCGCTGGCGCGCGCGCGCGCGCGCGCGCG	1077
DB	1157	ACATTAGAAGAAATGATGACAGCATGTCAAGGAGTGGGAGGAGCTAGCCACAAAGACA	1216
QY	1078	GTGCTGGCGGAGCGCATGAGCCAGGCGCAACA---CCAGGCTGATGATGATGATGATGATG	1134
DB	1217	GTGTTGGCTGAGGCAATGAGCCAAACAAATAGTTAAACATAGTATGATGATGATGATG	1276
QY	1135	TTCAAGGGCGCGCGCGCGCTGCTCAAGTCTTCAACTGCGGCAAGGAGGCGCGCGCGCG	1194
DB	1277	TTTAAAGGAATTAAGAAATGCTTAATGTTTAACTGTGGTAAAGAGGCGCATAGCC	1336
QY	1195	CGCACTGCG	1254
DB	1337	AGAAATTCAGGCG	1396
QY	1255	ATGAGGAGTGCAGCGGCG	1314
DB	1397	ATGAAAGACTGTACTGAGAGGCGCGCTAAITTTTAGGGAATAATTTGGCTTTCCCAAG	1456
QY	1315	GGCG	1374
DB	1457	GGAAGCGCAGGCAATTCCTTCAGAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1516
QY	1375	TTCCGCTTCGAGGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1434
DB	1517	TTGAGGTTTCGAGGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1576
QY	1435	ACAGCGCTGAACAGCGCTGTTCCGGAAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCG	1479
DB	1577	ACTTCCCTCAATCACTCTTTTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1621

## RESULT 7

US-09-184-418C-38  
 : Sequence 38, Application US/09184418C  
 : Patent No. 6492110  
 : GENERAL INFORMATION:  
 : APPLICANT: Hahn, Beatrice  
 : APPLICANT: Gao, Feng  
 : APPLICANT: Shaw, George  
 : TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN  
 : FILE REFERENCE: D6287  
 : CURRENT APPLICATION NUMBER: US/09/184,418C  
 : CURRENT FILING DATE: 1999-11-02  
 : NUMBER OF SEQ ID NOS: 112  
 : SEQ ID NO 38

```
/ LENGTH: 1486
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate-92RW009; gene-gag
US-09-184-418C-38

Query Match      48.1%; Score 710.8; DB 4; Length 1486;
Best Local Similarity 68.8%; Pred. No. 3.4e-99;
Matches 1022; Conservative 0; Mismatches 457; Indels 7; Gaps 3;

QY 1 ATGGGCGCCGCGCCAGCATCTTGGCGGCGGCAAGCTGAGAGGCTGGAGGCGCA-TCGG 59
DB 1 ATGGGTCGAGAGCGCTCAATATTAGAGCGCGGAAATTAGATGCTGGGAAAAAATTAG 60

QY 60 CCGCGCGCCGCGCGCAAGTGTACATGATGAGGACCTGGTGTGGCGCCAGCGGA 119
DB 60 CCGCGCGCCGCGCGCAAGTGTACATGATGAGGACCTGGTGTGGCGCCAGCGCG 119

QY 61 TTAAAGCGGGGGGAAAGAACATATATATGAAACACCTAGTAGTGGCAAGGGA 120
DB 61 TTAAAGCGGGGGGAAAGAACATATATGAAACACCTAGTAGTGGCAAGGGA 120

QY 120 GCTGGAGAAGTTTCGCGCTGAACCGCGGCTGTGGAGACGCGAGGCTGCAAGCAGAT 179
DB 120 GCTGGAGAAGTTTCGCGCTGAACCGCGGCTGTGGAGACGCGAGGCTGCAAGCAGAT 179

QY 121 GCTGGAGAAGTTTCGCGCTGAACCGCGGCTGTGGAGACGCGAGGCTGCAAGCAGAT 180
DB 121 GCTGGAGAAGTTTCGCGCTGAACCGCGGCTGTGGAGACGCGAGGCTGCAAGCAGAT 180

QY 180 CATCGCGAGCTGCACCGCGCGCTGCAGACCGCGAGGAGAGCTGAGACAGCTTTCAA 239
DB 180 CATCGCGAGCTGCACCGCGCGCTGCAGACCGCGAGGAGAGCTGAGACAGCTTTCAA 239

QY 181 AATGACACAGCTGCCAACAGCTCTTCAGACAGGAACAGATGAACTTAGCTATTATAA 240
DB 181 AATGACACAGCTGCCAACAGCTCTTCAGACAGGAACAGATGAACTTAGCTATTATAA 240

QY 240 CACCGTGGCCACCGCTGTACTGTGTCAGAGAGATGAGGTCGCCGCGACACAGGAG 299
DB 240 CACCGTGGCCACCGCTGTACTGTGTCAGAGAGATGAGGTCGCCGCGACACAGGAG 299

QY 241 TACAGTAGCAGCCCTTATTGTGTACATCAAAAGATAGATGTAAGACACACCGAGG 300
DB 241 TACAGTAGCAGCCCTTATTGTGTACATCAAAAGATAGATGTAAGACACACCGAGG 300

QY 300 CCGTGCACAGATCGAGGAGGAGCAGAACAGTGCAGCAGAGAAATCCACAGATGTGAG 359
DB 300 CCGTGCACAGATCGAGGAGGAGCAGAACAGTGCAGCAGAGAAATCCACAGATGTGAG 359

QY 301 CTTAGACAAGATAGAGGAGAACAAACAAAGTCAAGAAACACACAGCGCGAGACG 360
DB 301 CTTAGACAAGATAGAGGAGAACAAACAAAGTCAAGAAACACACAGCGCGAGACG 360

QY 360 CGCGACAGAGGCAAGGTGAGCGCAACTACCCCATCGTGCAGACACCTGCGAGGAGCAGAT 419
DB 360 CGCGACAGAGGCAAGGTGAGCGCAACTACCCCATCGTGCAGACACCTGCGAGGAGCAGAT 419

QY 361 AGCTGACAAAGGAAAGTCAGTCAAAATTAACCTATAGTCAAAATGACAAAGGCAAT 420
DB 361 AGCTGACAAAGGAAAGTCAGTCAAAATTAACCTATAGTCAAAATGACAAAGGCAAT 420

QY 420 GTTGACACAGGCGCATCAGCGCCGCGACCGTGAAGCGCTGGGTCAAGGTGATGAGAGAA 479
DB 420 GTTGACACAGGCGCATCAGCGCCGCGACCGTGAAGCGCTGGGTCAAGGTGATGAGAGAA 479

QY 421 GTTACACAGGCGCATATACCTAGAACCTTTGAATCGTGGGTAAAGTAATAGAGGAA 480
DB 421 GTTACACAGGCGCATATACCTAGAACCTTTGAATCGTGGGTAAAGTAATAGAGGAA 480

QY 480 GCGCTTACGCCCGAGGTGATCCCAIGTCCAGCGCCCTGAGCGAGGCGCGCCACCGCA 539
DB 480 GCGCTTACGCCCGAGGTGATCCCAIGTCCAGCGCCCTGAGCGAGGCGCGCCACCGCA 539

QY 481 GCGCTTACGCCAGAGGTAAACCCATGTTTACAGCATTAICAGAAGGAGCCACCGACA 540
DB 481 GCGCTTACGCCAGAGGTAAACCCATGTTTACAGCATTAICAGAAGGAGCCACCGACA 540

QY 540 GGACCTGAACAGCATGTTGAACACCGTGGCGGCGCCACCGCGCGCATGCGAGTCTGAA 599
DB 540 GGACCTGAACAGCATGTTGAACACCGTGGCGGCGCCACCGCGCGCATGCGAGTCTGAA 599

QY 541 AGATTTTAAACACCATGCTAAATACAGTGGGGGAGCATCAAGCAGCGCATGCAAAATGCTAAA 600
DB 541 AGATTTTAAACACCATGCTAAATACAGTGGGGGAGCATCAAGCAGCGCATGCAAAATGCTAAA 600

QY 600 GGACACCATCAACGAGGAGGCGCGAGTGGAGCGCGTGCACCGCTGCAGCGCGCGCG 659
DB 600 GGACACCATCAACGAGGAGGCGCGAGTGGAGCGCGTGCACCGCTGCAGCGCGCGCG 659

QY 601 AGATACAAATCATGAGGAGGCTGCAGGTGGATAGGTTACATCCAGTGCAGGAGGCGC 660
DB 601 AGATACAAATCATGAGGAGGCTGCAGGTGGATAGGTTACATCCAGTGCAGGAGGCGC 660

QY 660 CATCGCGCGCGCGAGATGCGCGAGCGCGCGCGAGCGACATCGCGGCGACACCGACAC 719
DB 660 CATCGCGCGCGCGAGATGCGCGAGCGCGCGCGAGCGACATCGCGGCGACACCGACAC 719

QY 661 TGTGTGCGCGCGAGTCAAGAGACCAAGGGGAGTGCATAGCAGGAACCTACTAGTAG 720
DB 661 TGTGTGCGCGCGAGTCAAGAGACCAAGGGGAGTGCATAGCAGGAACCTACTAGTAG 720

QY 720 CCGTGAAGGAGCAGATCGCTGTGATGACACAGCAACCGCGCGCGCGCGCGAGCATGTA 779
DB 720 CCGTGAAGGAGCAGATCGCTGTGATGACACAGCAACCGCGCGCGCGCGCGAGCATGTA 779

QY 721 CCGTTCAGGAACAAATAGCATGGATGACAAATTAACCCGCTATTCAGTGGGAGAAATTA 780
DB 721 CCGTTCAGGAACAAATAGCATGGATGACAAATTAACCCGCTATTCAGTGGGAGAAATTA 780

QY 780 CAAGCGGTGGATCATCTCGCGGCTGAACAAGATCGTGGGATGTACAGCGCGCGCTGAGAT 839
DB 780 CAAGCGGTGGATCATCTCGCGGCTGAACAAGATCGTGGGATGTACAGCGCGCGCTGAGAT 839

QY 781 TAAAGATGATAATCTCGGGTTAAATAAATAGTAAGATGATAGCGCTGTGAGAT 840
DB 781 TAAAGATGATAATCTCGGGTTAAATAAATAGTAAGATGATAGCGCTGTGAGAT 840

QY 840 CCGTGAACATCAAGCGCGCGCGCGAGCGCGCTTCGCCGACTACGTGAGCGGCTTCTCAA 899
DB 840 CCGTGAACATCAAGCGCGCGCGCGAGCGCGCTTCGCCGACTACGTGAGCGGCTTCTCAA 899

QY 841 ATTGGACATAAACAAGGCGCGCAAGGAGACCTTTTAGAGACATGCTAGACCGGTTCTTAA 900
DB 841 ATTGGACATAAACAAGGCGCGCAAGGAGACCTTTTAGAGACATGCTAGACCGGTTCTTAA 900

QY 900 GACCCTGCGCGCGGACGACAGCACCGCAGGAGGTGAGAACTGATGACGAGACCGCGCT 959
DB 900 GACCCTGCGCGCGGACGACAGCACCGCAGGAGGTGAGAACTGATGACGAGACCGCGCT 959
```

## RESULT 8

```
US-09-184-418C-4
; Sequence 4, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hann, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09184.418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 4
; LENGTH: 8992
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate-92RW009; 139.1624.gag; 1690.4428.pol(N-terminus)
; OTHER INFORMATION: 4373.4951.vif; 4891.5181.vpr; 5162.7801.tat; 5301.7958.rev;
; OTHER INFORMATION: 5403.5648.vpu; 5566.8148.env; 8150.8773.nef
US-09-184-418C-4
```

```
Query Match      48.1%; Score 710.8; DB 4; Length 8992;
Best Local Similarity 68.8%; Pred. No. 3.4e-99;
Matches 1022; Conservative 0; Mismatches 457; Indels 7; Gaps 3;

QY 1 ATGGGCGCCGCGCCAGCATCTTGGCGGCGGCAAGCTGAGAGGCTGGAGGCGCA-TCGG 59
DB 1 ATGGGTCGAGAGCGCTCAATATTAGAGCGCGGAAATTAGATGCTGGGAAAAAATTAG 198

QY 60 CCGCGCGCCGCGCGCAAGTGTCTACATGATGAAACACCTGCTGTGGCGCGAGCGCA 119
DB 60 CCGCGCGCCGCGCGCAAGTGTCTACATGATGAAACACCTGCTGTGGCGCGAGCGCA 119
```

199	TTAAAGCCAGGAGGGGAAAGAAAACAATATATATATGAAGAACACCTAGTATGGCAGAGTACGGA	258
120	GCTGGAGAAATTTCCGCTTGAAAGCCCGCCGCTGCTGGAGACACAGTCAGAGGCTGTCAATCAUAT	179
253	GCTGGAAGATTTGCACCTTAAACCTGACCTTTTAGAGACACAGAGAGCTGTANACAAAT	318
180	CATCGCCCATGTGACGCGCCGCTGGAGACCGGACCGGAGGAGCTGTAGAGGCTGTGTANA	239
319	AATGAGACAGCTGCACACAGCTCTTTCAGACAGAACAGATGAACCTTAGTGTCTATTATATAA	378
240	CACGGTGGCCACCTGTACTGGCTGCACGAGAAGATCGAGGTCGCGGACACCAAGGAGGC	299
379	TACAGTAGCAACCTCTATGTGTGTACATTCANAGATAGATGTANAGACACCCAGAGGC	438
300	CTGGACACAGATCGAGGAGGAGAGACAACAGTGCAGACAGAGATCCAGAGGCGGAGGC	359
439	CTTAGACAAGATAGAGAGACAACAAAACAAAAGTCAGCAAAAACACAGCAGGACAGGC	498
360	CGCGCAACAGGGCAAGTGTAGCCAGAACTACCCATCTGTGCAGAACTTCAGAGGCCAGAT	419
499	AGCTGACAAAGGAAAGTCACTCAAAATTACCTATAGTGCANAATGCACAGGGCAAT	558
420	GGTGCACACAGGCATCAGCCCCCGCACCTGTAAAGCCTGGGTGAGGTGATCGAGAGAA	479
559	GGTACACACAGGCATATCACTAGAACTTTGAATGCTGGTAAAAAGTAAATAGAGAGAA	618
480	GGCCTTCAGCCCGAGTGTATCCCATGTTCACCGCCCTGAGGAGGCGGCACACCCCA	539
619	GGCTTTAGCCAGAGATATACCCATGTTACAGCATATCAGAGGAGCCACCCACACA	678
540	GGACCTGAACACAGATGTTGAACACCTGGGCGGCCACACAGGCCCATCGAGATGCTGAA	599
679	AGATTTAAACACCATGCTAAATACAGTGGGGGACATCAAGCAGGCCATGCAAAATGCTAAA	738
600	GGACACATCAACGAGAGGCGCGCAGTGGGACCGGTGCACCCGTCGACCGCCGAC	659
739	AGTACATCATATGAGAGAGCTGCAGAGTGGGATAGGTACATCCAGTCGAGCGACGGCC	798
660	CATGCGCCCGCGGCAGATCGCGAGGCGCGCGGACGAGCATCGCGGGCACCCACCGAC	719
799	TGTTGGCCAGGCGCATATAGAGAACCAAGGGGAAGTGACATAGCAGAACTACTAGTAC	858
720	CTTGCAAGGACAGATGCGCTGGATGACCAACACCCCGCCATCCCGTGGCGCACATCTA	779
859	CTTTCAGGAACAATAGCATGGATGACAAATACCCACCTATTCCATGGGAGAAATTTA	918
780	CAAGCGGTGGATCATCTGGCGCTGAACAGATCGTGGGATGTACAGCGCCCGTGAGCAT	839
919	TAAAGATGATATTTCTGGGGTTAAATAAAATAGTAAGAAATCTATAGCCCTGTCAAGAT	978
840	CTTGACATCAAGCAGGCGCCAGAGAGCCTTTCGGGACTACGTGGACGGCTGTCTCAA	899
979	ATTGGACATTAACACAGGGCGCAAGGAGACCTTTTAGAGACTATGTAGACGGTGTCTTAA	1038
900	GACCTCTGCGCGCGAGCAGAGACCCAGGAGGTGAGAACTGGATAGCCGACACCTGTCT	959
1039	AACCTTAAGAGCGCAACAACTTCACAAGATGTAAAAATTTGGATGACACATACCTTGT	1098
960	GGTGACAGACCGCAACCCCACTGCAGAGACCTCTCGCGGCTCTCGGCGCCCGGCGGAC	1019
1099	AGTCCAAAATGCGAAGCCAGATTTGTAGAGCATTTTAAAGAGCATTAAGGCGCAGGCTTC	1158
1020	CTTGGAGGAGATGATGACCCCTTGCCAGGCGCTGGGCGGCCCGACGACACAGGCGGCT	1079
1159	ATTAGAGAANAATGATGACGATGCCAGGAGGTGGGAGGACCCCGGCGCTAAAGACAGGGT	1218
1080	GCTGGCGGAGGCGATGAGCCAGG---CCAAACACAGCGCTGATGATCAGAGAGGCAACTT	1138
1219	TTTGGCTGAGCAATGAGCCCAAGTACANCAACCAACATATATGATCGAGAGAGCATTT	1278
1137	CAAGGGCCCCGGCGCATCTGCTCAAGTGTCTCAACTGGCGCAAGAGGGCCACATGCGCGC	1196
1279	TAAAGGCGCAGAGAGATTTAATAGTGTTCAACTGTGGCAAGAGGAGCACCTTAGCCAG	1338

## RESULT 9

```

US-09-184-418C-82
: Sequence 82, Application US/091844418C
: Patent No. 6492110
: GENERAL INFORMATION:
: APPLICANT: Hahn, Beatrice
: APPLICANT: Gao, Feng
: APPLICANT: Shaw, Georgc
: TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
: TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
: FILE REFERENCE: D6287
: CURRENT APPLICATION NUMBER: US/09/184,418C
: CURRENT FILING DATE: 1999-11-02
: NUMBER OF SEQ ID NOS: 112
: SEQ ID NO 82
: LENGTH: 1496
: TYPE: DNA
: ORGANISM: Human immunodeficiency virus type 1
: FEATURE:
: OTHER INFORMATION: isolate=962M751.3; gene=gag
US-09-184-418C-82

```

```
Query Match      47.4%; Score 701.2; DB 4; Length 1496;
Best Local Similarity 68.8%; Pred No. 9.7e-98;
Matches 1032; Conservative 0; Mismatches 443; Index 25; Gaps 4;
```

Cy	1	ATGGGCGCCGCCGCCGATCTCTCGCGGGCGGCAAGCTGGACGCTCGGACGGCATCCGC	50
D6	1	AATGGTTCGAGAGCGCTCAATATTAAAGAGCGGGAATATTAGATGAATGSGAAAGAAATTAGG	60
Cy	61	CTCGCCCGCGCGCAAGAGTGTACATGATGAAGCACCTGGTGTGSGCCAGTCGGGAG	120
D6	61	TTAAGGCGAGGGGAAAAGACATATATGTAACACTTAAATATGGCAAGCAGGGAG	120
Cy	121	CTGGAGAAGTTGCGCTCTGAACCCCGGCGCTGTGAGACCAAGCGAGGCTGCAACGACATC	180
D6	121	CTGGAAGAATTGCGACTTAACTCTGGCTTTTAGAGACATCAGAAGCTGTAAACAAATA	180
Cy	181	ATCGGCGAGCTGCACCCGGCCCTGCAGACCGGCGAGCGAGCTGGAAGAGCTGTTCACAC	240
D6	181	ATACAACAGCTACACCAAGCTCTCCAGACAGGAACAGAGGAACTTAGGTCTATTATAAT	240
Cy	241	ACCGTGGCACCCCTGTACTGGTTCAGCAGAGAAGATCGAGGTCCGCGACACCAAGGAGGC	300
D6	241	ACAGTAGCACTCTCTATTGTGTACATGAAGAAGATAAAGGTACCAGACACCAAGGAACCC	300
Cy	301	CTGCACAGATCTGAGGAGGACGACACAAGTGCACAGCAGAGATCCACGACGGCCGAGGCC	360
D6	301	CTAGCAAGAATAGAGGAGAAACAACAAAAGTCAACAAAATAATACAAAAAACAGAGCG	360
Cy	361	GCCACAAAGGCGAAGGTGAGCGCAAACTACCCCATCTGTGCAAGACCTGCAGGGCCAGATG	420

361	Db	ACTGGC---GSAAGSTCACTCAAAATATTCCTATATGTCAGAAATTCCTCAAGGGCAATG	417
421	Qy	GTGCACGAGCCATACGCGCCGACCCCTGAACCGCTGGGGAAGGTGATCGAGGAGAAG	480
418	Db	GTACACGAGCTATATCATCTAGACCTTTTGAATGCAATGGGTAAAGATATAGAGAGCAAG	477
481	Qy	GCCTTCAGCCCGCAGAGTGATCCCATGTTCCAGCGCCCTGACGAGGGCGGACGCCAG	540
478	Db	GGTTTCAACCCAGAGGTAAATACCATGTTTACAGCATTTACAGAGGAGCCACCCACAA	517
541	Qy	GACCTGAACACGATGTTGAACACCGTGGCGGGCCACACGCGCCCATGCAAGATGCTGAAG	600
538	Db	GATCTAAACACCATGTTAATATACAGTGGGGGACATCAGACGCATGCAAA-TGTTAAA	597
601	Qy	GACACCATCAACGAGGAGCGCGGAGTGAGACCGGAGTGCAACCCCTGCACGCGGCCCG	660
598	Db	GATACCATCAATGAGGAAGCTGACAGATGGGATAGGTACATCAGGTACATGCGAGGCT	617
661	Qy	ATCGCCCGCGCCAGATCGCGAGCCCTGCGGAGAGGACATCGCGGACATCAGCAGCAC	720
658	Db	ATTGCACCGCCAAATAAGAGNACCAAGGGGAAGTCNCATAGCAGAACTACTGATACC	717
721	Qy	CTGAGGAGCAGATCGCTGGATGACACGACCAACCCCGCATCCCGGTGGGCGACATCTAC	780
718	Db	CTTCAGGACAAATAGCATGATGACAAATATACCCACCTATTCCAGTGGGAGACATCTAT	777
781	Qy	AAGGGGTGGATCATCCCTGGCGCTGAACAAAGATCGTGGGATGTACAGCCGCTGACATC	840
778	Db	AAAAGATGGGATATCTGGGGTAAATATAAATATAGAAATGTACAGCTCTGACGATTT	837
841	Qy	CTGCATCATAGAGGGCGCCCAAGAGACGCTCTCGGCATCTAGTGGACGCTCTCTCAAG	900
838	Db	CTGACATTAACAGGACCAAGGACCCCTTAGGCACTATGTAGATCGTCTCTTAAA	837
901	Qy	ACCTGCGCGCCGAGAGACACCCAGGAGTGAAGAAGCTGGATGACCGACACCTCTCTG	960
898	Db	ACTTTAAGAGCTGAACAGGTACACAAAGTGT-AAAAATTGGATGACAGACGCTCTTG	955
961	Qy	GTGCAGAACCCACCCCGACTCGACACCATCTCGCGGCTCTCGCCCGCGGGCGAGC	1020
957	Db	GTTCAAATCGCAACCCAGATGTGAGACCATTTAAGGGCATTTAGACCCAGGGGCTACA	1015
1021	Qy	CTGAGGAGATGATGACCGCTCGCAGGGGCTGGCGCGCCCGACCCACAAAGCCCGGTG	1080
1017	Db	TTAGACAAATGATGACACCATGTACAGGAGTGGGGGACCTGCCACAAAGCAAGATTT	1076
1081	Qy	CTGGCCAGCGGATGAGCCAGG---CCACACACAGCTGATGTGASAGAGCAACTTC	1137
1077	Db	TTGGCTGAAGCAATGAGCCAAAGTAAACAATACAAACATAATGATGCAGAAAGCAATTT	1136
1138	Qy	AAGGGCCCGCGGCATGTCTCAAGTGTCTCAACTGGCGGAGGAGGCCACATCGCCCGC	1137
1137	Db	AAAGGCCCTTAAGAATTTGTAAAGTGTTCACCTGTGGCAGGAGGCGCATATAGCCAGG	1136
1198	Qy	AACTGCGCGCCCGCCCGCCAGAGAGGGCTGCTGCAAGTGGCGCAAGAGGCGCCACGATG	1257
1197	Db	AATTTGACGGCTCTCTGGGAAAAAGGCTGTGGAAATGTGGAAAGGAGGACACCAATG	1256
1258	Qy	AAGACTTGCACCGAGCGCCAGGCCAACTTCTTGGGCAAGATCTGGGCCCGCCACAGGCG	1317
1257	Db	AAAGACTGTACTGAGAGACAGGCTTAATTTTATAGGAAAAATTGGCCCTTCCCAAGAGCG	1316
1318	Qy	CGCCCCGGCAACTTCCTTCGACAGCGCGCGGAGCGCCAGCCCGCC-----	1359
1317	Db	AGGCTGGGGAACITCTCTTACAGACAGACAGAGCCACAGCCCGCCACCATGCTCCACAGCC	1376
1360	Qy	CCCCCGCGCGAGAGCTTCGCTTCGAGGAGACCAACCCCGCGCGCCAGAGAGAGCAAG	1419
1377	Db	CCACCAGTAGAGAGCTTCAGGTTGAGGAGACACCCCTGCCCCGAGCGCAGGACAGAAA	1416
1420	Qy	GACCGCGAGACCTTGACACGCTGAAGAGGCTGTTGCGGCAACGACGCCCCCTGAGGCCATTA	1479

```

Db      : 437  GACAAGGACCCCTAACTGCCCCTCAAAATCACTCTTTGGGACGCCAGCCCTTGTTCTCAATAA 1499

RESULT 10
US-09-184-418C-9
: Sequence 9, Application US/09184418C
: Patent NO. 6492110
: GENERAL INFORMATION:
: APPLICANT: Hahn, Beatrice
: APPLICANT: Gao, Feng
: APPLICANT: Shaw, George
: TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
: TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
: FILE REFERENCE: D6287
: CURRENT APPLICATION NUMBER: US/09/184.418C
: CURRENT FILING DATE: 1999-11-02
: NUMBER OF SEQ ID NOS: 112
: SEQ ID NO 9
: LENGTH: 8972
: TYPE: DNA
: ORGANISM: Human immunodeficiency virus type 1
: FEATURES:
: OTHER INFORMATION: isolate=962M751.3; 137.1632:999; 1419.4475:pcl;
: OTHER INFORMATION: 4380.7938:wif; 4898.3188:vpr; 5169.7814:lat;
: OTHER INFORMATION: 5308.7938:rev; 5407.5657:vpu; 5585.8128:env;
: OTHER INFORMATION: 8130.8753:nef
US-03-184-418C-9

Query Match      47.48; Score 701.2; DB 4; Length 8972;
Best Local Similarity 66.84; Pred. No. 9.3e-98;
Matches 1032; Conservative 0; Mismatches 443; Indels 25; Gaps 4;

Oy      1  ATGGCCCGCCGCGCAGCATCTCTGCGGCGCGGCAAGCTGGACGCGCTCGGAGCGCATCCGC 60
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      137  ATGGTGGGAGGCGTCAATATTAGAGGCGGGAANTTAGTGAATGGGAAGNATTAGG 196

Oy      61  CTGGCCCCCGCGCGCAAGAGTGTACATGATGAAGCAACCTGTTGGTGGGCCACGCCGCGAG 120
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      197  TTAAGGCGCGGGGCAAAAAGACATATATGATGAACACTTAATATGGCAAGCAGCGAG 256

Oy      121  CTGAGAAGTTGCGCTGACCCCGGCTCTGTGGAGACCGAGGAGCTGCAAGCAGATC 180
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      257  CTGAAAAGATTGCACTTAACCCCTGGCCTTTTAGAGACATCAGAAGCTGTAACAAATA 316

Oy      181  ATCGCCAGCTGCACCCCGCCTCGACCGCGGAGGAGCTGGAAGAGCGCTGTTCAC 240
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      317  ATACAAGCTACACCAGCTCTCCAGCAGGACACAGGNACTTAGTTCATTATATAT 376

Oy      241  ACCGTGTCACCCCTGTACTGTGTCACGAGAAGATCGAGGTTCGCGACACCAAGAGGCC 300
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      377  ACAGTAGCACTCTCTATTGTGTACATGAAAAGATAAAGGTACGAGACACCAAGGAAGCC 436

Oy      301  CTGACAAAGATCGAGGAGGACCAACANGTGCACGACAGAGATCCAGCAGGCGGAGGCC 360
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      437  CTGACAAGATAGAGGAGAACAAACAAAGTCAACAAAATAATACAAAACACAGAAGCG 496

Oy      361  GCCGACAAGGCGAAGGTGAGCCGAAGCTACCCCATCGTCAGAACCTCCAGGCGCAGAT 420
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      497  ACTGCC---GGAAAGGTCAGTCAAAATTATCCTATATGTCAGAAATCTCCAGGCGCAATG 553

Oy      421  GTGCACAGGCCATCAGCCCGCCGACCCCTGAACCCCTGGTGAAGTGTATCGAGGAGAAG 480
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      554  GTACACAGGCTATATCATCATAGAACTTTGAATGCATGGGTAAAGTAATAGAGGAAG 613

Oy      481  GCCTTAGCCCCGAGGTGATCCCATCTGTACCGCCCTGAGCGAGGCGGCCACCCGCCAG 540
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      614  GGTTTCAACCCAGAGGTAATACCCATGTTTACAGCATATTCAGAAAGAGCGCCCCCAAA 673

Oy      541  GACCTGAACACGATGTTGAACACCGCTGGCGGCCACCAAGGCGGCTGTCAGATGCTGAAG 600
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      674  GATCTAACACCATGTTTAATACAGTGGGGGACATCAAGCAGCCATGCAATGTTAAA 733

Oy      601  GACACCATCAACGAGGAGGCGCGGAGTGGATCGCGTGCACCCCGTGCACGCGCGGCC 660

```





```
QY 1075 CGCGTGGTGGCGGAGGATGAGCGAGG-----CGAACACAGCGGTGATGATGGA 1125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1217 AGAGTATTGGCTGAGGCAATGAGGCAAGTAGAACAAATACAAATACAGCCATAATGATGCA 1276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1126 AAGAGCAATTCAGAGGCGCGCGGCGCATCGTCAAGTGTTCACACTGCGGCAAGAGGCT 1185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1277 AAGGCAACATTAGGCGCAAGAAATTTGTTAAATGCTCAACTGTGCGCAAGAGGGA 1336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1186 CACATCCCGCGCACTCCCGCGCGCGCGCGCAAGAGGCGTCTGGAAGTCCGCGAGAGAG 1245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1337 CACATGCCAGAAATTCAGAGGCGCGCTAGGAAAAGGGCTGTGGAATATGCGAAGAGAA 1396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1246 GGCCACACAGATGAAGGACTGCACGAGCGCGCACTTCCTGGGCAAGATCTGGCGC 1305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1397 GGACATCAGATGAAGACTCCACAGAGACAGAGCTAA:TTTTTAGGAAAT:GGGCT 1456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1306 AGCCACAAGGCGCGCGCGCGCACTTCTCTCAGAGGCGCGCGCGAGCCACCGCGCGCG 1365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1457 TCCAGCAAGAGGAGCGCGAGGAAATTTCTCCAGAGCGAGGCGAGAACAGCGCGCGCA 1516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1366 GCGGAGAGCTTCGCGTTC:---GAGGAGACCAACCGCGCGCGCGAGAGCAGACCAAGAG 1422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1517 GCAGAGAGCTTCGGGTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1423 CGGAG-----ACCGTGAACAGCGCTGAGAGAGCGCTTCGCGCAACGACCGCGCGAG 1476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1577 CACAAGGAACCTCCCTTAGCTTCCTCAGATCACTCTTTGGCAGCGACCCCTTGTACAG 1636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-09-184-418C-12
; Sequence 12, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONING AND SEQUENCES FOR NON-SURTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184.418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 12
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate-94HR020; gene-gag
US-09-184-418C-12

Query Match 44.5%; Score 658.8; DB 4; Length 1479;
Best Local Similarity 67.9%; Pred. No. 2.3e-91;
Matches 969; Conservative 0; Mismatches 447; Indels 12; Gaps 3;

QY 1 ATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGGGTCGAGAGCGTCAGCTATTAAAGCGGCGGAAATTTAGATCTTGGGCAAGAGAGAG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 TTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 CTGAGCAAGTTCGCGCTGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 CTAGAACGATTTCGACTTGTATCCAGGCGCTTCTAGAACATCAGAGAGGCTTCGAAATA 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 ATCGCGCAGCTGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 A*AGCAGAGTTACAAACCATCCCTTCAGACAGATCAGAAGAGCTCAAAATCATATATAAT 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 ACCGTGGCGACCCCTGTACIGGTCACACAGAGAGATCGAGGTCCCGGACACCAAGAGGCC 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 ACAATAGCAGCTCTCTATTATGTATACATAAAAGTAGAGGTAAAGACACCAAGAGGCT 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 CTGGCAAGATCGAGGAGGAGCAGAAACAAGTGCACGACAGAAGATCCACGACGCGCGAGGCC 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TTAGAGAGAGCTAGAGGAGAGNACAAACAAGGTCCGCAAAAGACACACAGACGAGCTGCT 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 GCCGACAAGGCGCAAGGTGAGGCCAGAACTACCCCATCTGTGCAGAACTCCAGGCGCGAGAT 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 GAAAAAG-----GCTCAGTCANAATTAACCTATAGTACAGAACTTTCAGGCGCAAAATG 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 GTCCACAGGCGCATCAGCCCGCGCACCTGAACGCTGGGTGAAGGTGATCCAGGAGAG 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 415 GTACACCGAGCTTTTAATCAGCTAGAACTTTAATGATGGGTAAAGGTGATAGAGAGAG 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 GCCTTCAGCGCGGAGGTGATCCCATGTTCACCGCGCTGAGCGAGGCGCGCACCCCGCAG 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 475 GCTTTTAGTCCAGAAGTAATACCCATGTTTCAGCATTAICAGAAGGGCGCCACTCCACAA 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 541 GACCTGAACACAGATGTTGAACACCGTGGCGCGCCACCGAGCGCGCATCAGATGCTGAAG 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 535 GATTTAAACACCATGTTAAATACAGTGGGGGACATCAAGCAGCGCATGCAAAATGTTAAA 594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 601 GACACCATCAAGCAGGAGGCGCGCGAGTGGGACCGCTGCACCCCGTGCACCGCGCGGCC 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 595 GACACCATCAATGAGGAGGCTGCAGAAATGGGACAGATTACATCCAAACACAGCGCAGGCC 654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 661 ATCGCGCGCGCGCAGATCGCGAGCGCGCGCGCGCGCGCATCGCGCGCACACCGAGCAC 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 655 ATCCCGCGCGCGCAGTAAAGGAACTTAGGGAACTGATATAGCTGGAATCTAGTACC 714
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 721 CTGAGGAGCAGATCGCTGTGATGACACAGCAACCGCCCATCTCCCGCTAGCTGGGCGACNTCTAC 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 715 CTTAGGAACAATAACAATGATGACAGCAACCGCCCTGTCCCACTGGGAGAAATGAT 774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 781 AAGCGTGGATCATCTCGGCTGAAACAAGATCGTGGCGATGTACAGCGCGCGTGAAGCATC 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 775 AAAGATGGATCATCTAGGATTAATAAATAGTAAGAACATTTTAAAGCATTTGGGACCGGCTACA 834
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 841 CTGGACATCAAGCAGCGCGCGCGCGCGCGCTTCGCGCTAGCTAGCTGGACCGTCTTCTCAAG 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 835 TGGACATAAGACAAAGCGCCAAACAACCGCTTACAGACTATGTAGCAGGTTCTTTAAA 894
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 901 ACCCTGCGCGCGCGCAGCAGCACCAGGAGGTGAGAAGTGGATGACCGACACCTGCTG 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 895 ACCCTAAGAGCTGAGCAAGCTACACAGGAAGTAAAGGGTTGGATGACAGACACCTTGTG 954
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 961 GTGCAAGCGCCCAACCGCGCGCGCGCGCGCGCTGCGCGCTCTCGCGCGCGCGCGCGCG 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 955 GTCCAAATGCCAACCGCGAGATTGAAGACCATTTTAAAGCATTTGGGACCGGCGCTACA 1014
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1021 CTGGAGGAGATGATACCGCGCTGCCAGGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCG 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1015 CTAGAGGAATGATCAGACGATGTCAAGGAGTGGGAGCGCTAGGCAATAAGGCAAGATT 1074
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1081 CTGGCGGAGCGCATGAGCGAGG---CCACACCGCGTGTATGATCCAGAAAGCAACTTC 1137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1075 TTGGCTGAGGCAATGAGCGCAAGCAAAATACAGCTAATAATGATGACAGAAAGTAACCTT 1134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1138 AAGGCGCGCGCGCGCATGCTCAAGTCTTCACTCGCGCAAGGAGGCGCGCATCGCGCGC 1197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1135 AAGGCGCAAGAGAGAAATTTTAATGCTTTAATGTGGCAAGAGAGACACATAGCCAAA 1194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1198 AACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1195 AATTGCGAGCGCGCGCTAGAAAAAAGGCTGTGGAGTGTGGAGAGAGGACACCAATG 1254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1258 AAGGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```





```
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/124,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 48
; LENGTH: 4307
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=92NS063; yone-qag-pol fusion
US-09-184-418C-48

Query Match          44.4%; Score 656.6; DB 4; Length 4307;
Best Local Similarity 66.9%; Pred No. 4.7e-91;
Matches 1005; Conservative 0; Mismatches 474; Indels 24; Gaps 4;

QY 1 ATGGGCGCCGCGCCAGCATCCCTGCGCGCGCGCAAGCTGGACGCTGGGACGCGCATCCGC 60
DB 1 ATGGGTGCGAGAGCGTCACTATTAAGCGGGGMAATTAAGATGATGGGAAAAATTCGG 60

QY 61 CTGCGCCCGCGCGCGCANGAGTGCTACATGATGAGCACCTGGTGTGGCCAGCGCGAG 120
DB 61 TTGAGCCCGCGCGCGCGCAAAAATATAGATGAACATTTAGTATGGGCAAGCAGGAA 120

QY 121 CTGGAGAGCTTCGCCCTGAACCCCGGCTGCTGGAGACACGAGGCTGCGACGAGATC 180
DB 121 CTGGAGAGATTTGCACATTACCCCTGACCTTTAGAAACAACAGAGGTTGTACCAAA 180

QY 181 ATCCGCCAGCTGCACCCCGCCCTGAGACCGCGCAGGAGAGCTGAGAGCCTGTTCAAC 240
DB 181 ATGACAGAGCTGCAACCATCTCTCCAGACAGGACAGAGGAGATTAAATCATTTAA 240

QY 241 ACGGTGGCCACCCCTGTACTGCGTGCACGAGAAGATCGAGTCCGCGCACCAAGAGGCC 300
DB 241 ACAGTAGCAACCCCTCTATTGTGTCATCAAGGATAGAGTTAAAGACACCAAGAGCT 300

QY 301 CTGGACAGATCGAGGAGGAGCAGAACAAAGTGCACGAGAGATGCCAGAGATCCAGCAGC 354
DB 301 CTGAGGAGAGTGGAAAAATCAAAAGAACACAGTCAAGAAAGAACAGAGAGGAGCAATG 360

QY 355 GAGGCGCGCGCAGGAGGAGGAGTGAACCAAGTACCCCATCTGTCAGAGACCTGCAGGCG 424
DB 361 GGTAAAGAAACAGCAGCCCAAGTATGCCAAATATCTCTATAGTGCAGATGTCACAGGG 420

QY 415 CAGATGGTGCACAGCGCATACGCCCGCGACCCCTGAACGCTGGGTGAAGTGAATGAG 474
DB 421 CAAGTGGTACACAGCCCATATCACTAGGACTTTAAATGCTGGGTAAAGAAATAGAA 480

QY 475 GAGAGGCTTCAGCGCGCGAGTGATCCCATGTTCACCGCCCTGAGCGAGGCGCCACC 524
DB 481 GAAAGAGCTTCAGTCCAGAGATATACCCATGTTACAGCATTAACAGAGAGGACACC 540

QY 535 CCCCAGGAGCTGAACACAGATGTTGAACACGCTGGCGCGCCACCAAGCGCGCATGCAAGT 594
DB 541 CCACAGAGTTTCAATACCATGTAACACCGCTGGGCGGCGCATCAAGCAGCTATGCAATG 600

QY 595 CTGAGGACACCATCAAGCAGGAGCGCGCGAGTGGGAGCCCGCTGCACCCCGTGCAGCC 654
DB 601 CTAAAGATTTCTATTATGAAGAGCTGCAGATGGGATAGGCTTACAACACAGCGCA 660

QY 655 GCGCCCATCGCCCGCGCGAGATGCGGAGCGCCCGCGAGCGACATCGCGCGGACCAACC 714
DB 661 GGACCTATTCCACAGCGCGAGATAGAGAACACAGGCGGAAGTATAGCAGAGACTACT 720

QY 715 AGCACCTCGCAGGAGCAGATCCCTGGATGACGAGCAGACCCCGCCATCCCGTGGGCGAC 774
DB 721 AGTACCTCGAGGAAACAAATACATGGAATGATGACCAACCAACCCACCTATCCCAAGTGG 780

QY 775 ATCTACAAGCGTGGATCATCTCGGCTGAACAGAGATCGTGGGATGTACAGCCCGTGC 834
DB 781 ATTTATAAGAGATGATATTTCTGGGTTTAAATTAATAGTGAAGTGTAGCCCTGTC 840

QY 835 AGCATCTGGACATCAAGCAGGCGCCCAAGGAGCCCTTCCGCGACTAGCTGAGCCGCTC 894
```

```
DB 841 AGTATTTTAGACATAAAACAAGGGCCAAAAGACCTTACAGAGATTATGTGGATAGGTT 900
QY 895 TTCAAGACCTTGGCGCGCGAGCAGACACCCAGAGAGTGAAGAACTTGGATACCCACACC 954
DB 901 TTAAAACTTTGAGAGCTGAGCAAGCCACACAGGAGTTAAAACTTGTATGACAGACACC 960

QY 955 CTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCTCTGCGGCTCTGCGCCCGGC 1014
DB 961 TTGTTGGTCCAAAATSCGAACCCAGATTGTAAAGACCATCTTAAGAGCATTAGGAGCAG 1020

QY 1015 GCGACGCTGAGAGAGATGATGACCCCTGCGCAGGCGCTGGCGCGCCGCCAGCAACAGGCC 1074
DB 1021 GGTACACTAGAGAATGTTGACACCACTGCAAGAGTGGGAGGACCCAGCCCAAGAGCA 1080

QY 1075 CCGCTGCTGGCGCGAGCGGATGAGCCAGGCCA-----ACACAGCGTGTATGATCAG 1125
DB 1081 AGAGTTTATGCTGAGGCAATGAGCGCGGCAACAGGTATACATCAGCAGGCCATATGATG 1140

QY 1126 AAGACCACTTCAAGCGCGCCCGCGGCATCGTCAAGTCTTCAACTGCGCAAGGAGGC 1185
DB 1141 AAAACAATTTTAAGCGCGCCCGGAGAGAGGTATTAAGTGTTCAACTGTGGCAAGGAAG 1200

QY 1186 CACATGCGCGCAACTGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAG 1245
DB 1201 CATCTAGCCAGAAATTCAGCGGCCCTTAGGAAAAAGGCTGTGSAATGTGGAAGGAG 1260

QY 1245 GGCACCCAGATGAAGAGACTGCACCAKCGCCAGGCCAACTTCCTGGGCAAGATCTGGCC 1305
DB 1261 GGACATCAATGAAAGACTGCACAGAGAGACAGGCTAATTTTTTAGGMAAATTTGCGCT 1320

QY 1306 AGCCACAGGGCGCGCGCAACTTCCTGCAAGACCGCCCGCGAGCCCGCCCGCCCGCC 1365
DB 1321 TCCACAGAGGCGCGCAGGGAATTTCTTCAGAACAGCCAGAGCCACAGCCCGCCACC 1380

QY 1366 GCGGAGAGCTTCGGCTTC---GAGGAGACACCCCGCGCCAGAGAGCAGAGAGCAAGGAC 1422
DB 1381 GCAGAGAGCTTCGGGTTCGGAGAGGAGATAGCCCTTCCTCTGAAGCAGGAGCGCGAGAA 1440

QY 1423 CCGCAGACCC-----TGACCACGCTGAAGAGCTGTTGCGCAACACGCCCTCAGGCAG 1476
DB 1441 AAGGAATCACCTCCATTAACTCCCTCAATCACTCTTTGGCAAGCGCCCTAGTCACAG 1500

QY 1477 TAA 1479
DB 1501 TAA 1503
```

Search completed: September 22, 2003, 22:53:41

Job time : 107.531 secs





```

Db 330 CATCGGCTTCTGTGTCGACGAGCTGGGCTCGACGCCGACCCCTGCAAGGTCTGTCATA 389
QY 504 CATGTTACCGCCCTGAGCGAGGCGCCACCCCGCAGGACCTTGAACGATGTTGAACAC 563
Db 390 CATCGAGCAGCACTGCCCGACATCGCGAGGGGTGCAGGCCACTTCAACAAAGCGCC 449
QY 564 COTGGCGGCCACCGAGCGCCCTCATGATGCTGAAGGACACCATCAACGAGGCGCGC 523
Db 450 CGAGGAGATCGCGCGCGGACCGAGGCGCCACATGTTGGCTAGCGCCACCGAGGACCC 509
QY 624 CGAGTGGGACCGGTGCACCCCGTGCACCGCGGCCCATCGGCCCGCGGCAGATCGCGGA 583
Db 510 CGAGCTGATGCGCCCTGACGACCTGCTCGCCACCAAGCTCGCGCGCCCTCACCGAGT 569
QY 684 GCGCCGCGCAGCATCGCGCGGACACACGACCGCTTGCAGGAGCAGATGCGCTGNT 743
Db 570 CCGCAGAAGACGCACTGCGCTTGGCTAGCGCCGCGCAGGACCCAGGTCAACGTTCA 629
QY 744 GACCAAGACCGCCCATCCCGTGGGAGACATCTACAGCGGTGATCTCTCGGCT 803
Db 630 GTACCTCAACGACGCGCGCGCCATGCTCCCGTCCGCTCCACACCGTCTCATCTCAC 689
QY 804 GAACAAGATCGTGGGATCTACAGCCCGTGGAGCATCTGGAGCATCAAGCAGGCGCCGA 863
Db 690 CCAGCAGCAGAGAGCTTACCAACGACGAGATCCCGCGCACCTCAAGGAGCAGTCTAT 749
QY 864 GGAGCGCTTCCCGGACTAGTGGACCGCTTCTTCAAGACCTT 905
Db 750 CAAGCGGTCATCCCGCACAGTACCTCGACGAGAGACCAT 791

```

## RESULT 3

```

CB662794
LOCUS
DEFINITION
OSJNEd07C09.f OSJNEd Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEd07C09 5', mRNA sequence.
CB662794
VERSION
KEYWORDS
SOURCE
ORGANISM

```

```

Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
; (bases 1 to 821)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Ming,K. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe oryzae
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu/
PCR Primers
FORWARD: gta aaa cga cgg cca atg
BACKWARD: gga aac agc tat gac cat g
Plate: 07 row: C column: 09
Seq primer: gta aaa cga cgg cca atg.
Location/Qualifiers
1..821
/oraniser="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEd07C09"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"

```

## FEATURES

```

source
RESULT 4
CB666192
LOCUS
DEFINITION
OSJNEd12P19.f OSJNEd Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEd12P19 5', mRNA sequence.
CB666192
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 824)

```

```

/cdona.lib="OSJNEd"
/note="vector: pBluescript II KS-; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (C9240-1)".
BASE COUNT 166 a 309 c 230 g 116 t
ORIGIN

```

```

Query Match 5.3% Score 78; DB 14; Length 821;
Best Local Similarity 44.4% Pred. No. 0.00074;
Matches 312; Conservative 0; Mismatches 390; Indels 0; Gaps 0;
QY 204 GCAGACCGGAGCGAGGAGTGAAGAGCCCTGTTCAACACCGTGGCCACCTGTACTGCGT 263
Db 88 GCAGATGAATGCGCGGAGAGCTTCCCTTCCCTCCGAGTCCGTAACGAGGTCA 147
QY 264 GCAGAGAAGATCAGTCCCGGACACCAAGAGGCGCTGCACAAGATCAGAGGAGCA 323
Db 148 CCGGAGAAGCTGTGCACCAGGTGTGGAGCGGTGCTCGAGCGTGCCTCGGCCAGGA 297
QY 324 GAACAAGTGGCAGCAGAGATCCAGCAGGCGCGAGGCGCGCAAGGSCAAGGTGAGCCA 383
Db 208 CCGGACAGCAAGGTGGCGTGGAGAGCTGCACCAAGCAACATGTTGTTGCG 267
QY 384 GAATACCCCTCTGTCAGAACCTGCAGGCGCAGATGTCACCGCCCATCAGCCCGC 443
Db 268 CGAGATCACCAAGCCACCGTGCAGTACGAGAAGATGTCGCGCACCTGCCCGG 327
QY 444 CACCTGAAGCGCTGGTGAAGGTGATCGAGGAGAGGCGCTTCAGCCCGCAGGTATCC 503
Db 328 CATCGGTTGTTGTCGACGACGTGGCGCTCGAGCGCGACCGTTCGANGTGTCTGTCAA 387
QY 504 CATGTTACCGCCCTGAGCGGCGCACCCCGCAGGAGCTGAACACGATGTTGAACAC 563
Db 398 CATCGACGAGCAGTCGCCCGACATCGCGCAGGGGTGCAGCGCCACTTCAACAGCGCC 447
QY 564 CGTGGCGGCGACCGCGCCATGAGATGCTGAAGGACACCATCAACGAGGAGGCGCG 623
Db 448 CGAGGAGATCGGCGCGCGGACCGACCGGCGCATGTTGGCTAGCGCCACCGACGAGCC 507
QY 624 CGAGTGGAGCGCTGCACCGCGTGCAGCGCGCGCCCATCGCCCGCGCAGATGCGCGA 683
Db 508 CGAGCTGATGCGCTCAGCCACGTCCTCGCCACCAAGCTCGGCCCGCCCTCACCGAGGT 557
QY 684 GCGCCGCGGAGCGACATCGCGCGCACCGACCGACCGCTTCAGGAGTATGCGCTGAT 743
Db 568 CCGCAAGAACGGCAGCTCGCGCTGGCTCAGCGCGCGCAAGACCCAGGTCAACGTTGA 627
QY 744 GACCAAGAACCCCGCATCGCGTGGCGACATCTACAGGCGGTGATCTCTGGCGCT 803
Db 628 GTACTCAACGAGCGCGCGCGCATGTTCCCGCTCCGCTCCACACCGCTCTCTCA 687
QY 804 GAACAAGATCGTGGGATGTACACCCCGCTGAGCATCTCGACATCAAGCAGGCGCCCAA 863
Db 688 CCAGCAGCAGAGACCGTCCACACGACGAGATCGCGCGGACCTCAAGGAGCAGTCTAT 747
QY 864 GGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAAGACCTT 905
Db 748 CAAGCGGTGATCCCGGACCAAGTACCTCGACGAGAGACCAT 789

```

AUTHORS	Jantanusriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,S., Wing,R. and Wang,X.
TITLE	Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe oryzae
JOURNAL	Unpublished
COMMENT	Contact: Rod Winq Arizona Genomics Institute University of Arizona Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA Tel: 520 626 3967 Fax: 520 621 9288 Email: http://genome.arizona.edu PCR Primers FORWARD: gta aaa cga cgg cca gta g BACKWARD: gga aac agc tat gac cat g Plate: 12 row: p column: 19 Seq primer: gta aaa cga cgg cca gta g. Location/Qualifiers 1. 524 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="mRNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /clone="OSJNEd2p19" /tissue_type="leaf" /dev_stage="3 week" /lab_host="DH10B" /clone_lib="OSJNEd" /note="vector: pBluescript II KS +; Site1: EcoRI; Site2: XhoI; 24 hrs after inoculation with Rice Blast (CG245-1)" BASE COUNT 166 a 312 c 232 g 114 t ORIGIN
Query Match	5.38; Score 78; Db 14; Length 824;
Best Local Similarity	44.4%; Pred. No. 0.00074;
Matches 312:	Conservative 0; Mismatches 390; Indels 0; Gaps 0;
QY	204 GCAGCCGGCAGCGAGGAGCTGAAGAGCCTGTTCACACCCGTGGCCACCCTTACTCTGGCT 263 
Db	98 GCAGATCAAAATGGCGCGGAGAGTGTTCTCTTTCACCTCCGAGTCGGTGAACGAGGGTCA 157 
QY	264 GCAGGAGAAGATCGAGGTCCGCACACCAAGAGAGGCCCTGGACACAGATCGAGGAGGACA 323 
Db	158 CCGGACAAGCTGTGGCACCATAGTGTGGACGGGGTGTCTGCAGCGGTGGCTTCGCCACGGA 217 
QY	324 GAACAAGTGCACAGCAAGATCCACAGGCGGAGCGCCGACAGGCGCAAGGTGAGGACA 383 
Db	218 CCCGACAGCAAGTGCGGTGCGAGAGCTGCACCAGNCCAACATGGTGNTGGTGTTCGG 277 
QY	384 GAATACCCGAATCTGCAGAACCTTCAGGGCGAGATGGTGCACAGGACATCAGGCGGG 443 
Db	278 CGAGATCACCAAGCGCACCGCTGCCTACTACGAGAAGATCTCCGCGACATCTGCCGGGG 337 
QY	444 CACCTGTAAACGCTTGGGTGAAGTGATCCGAGGAGAAGGCGCTTCAGCGCCGAGAGTGATGCC 503 
Db	338 CATCGCTCTGTGTCCGACAGCTCGGCTTCGACGCCGACGCTGCAGAGGTCTCTCAAA 397 
QY	504 CATGTTTCACGCCCTTGAGCGAGGGCCCCACCCCACAGGACCTTAACACAGATGTTTAACAC 563 
Db	358 CATCGACGACGATCGCGCCGACATCGCGCAGGGGGTGCACGGGCCATTTCACCAAGCGGCC 457 
QY	564 CGTGGCGGCCACAGCGGCCCATCGCATGCGATGTCGAGAGACACATCAAGGAGGAGGCCGC 623 
Db	458 CGAGGAGATCGGCGCCCGGCAACGAGGCCACATGTTTCGGCTACGGCACCGAGCAGACGCC 517 
QY	624 CGAGTGGGACCGCGTCACGCGCTGCACGCGGCCGCCCATCGCGCCGCGGCAGATGCGAA 683 
Db	518 CGAGCTGATGCGCCTCAGCAAGTGTCTGCGCCACCAAGCTCTCGCGCGCGCTCAGCGAGGT 577 
QY	684 GCCCGCGGCCAGGACATCGCGCGCACCACTACGACACCTTCGAGGAGACATAFGGCTGGAT 743 
Db	578 CCGAAGAACCGGCACTTCGCTCTGGTTCAGCGCCCGCAGCGCAAGACCCAGGTCACCGTTGA 637 

CY 744 GACACAGAACCCCGCCATCCCTGGGCGNACTTACAAGSNGTGGAATCATCTCTGGGCCT 803  
 |||| |  
 DB 618 ATACCTCAAGACGCCGGCGCCATGTGTCCCGTCGCGCTCCACACGGTCTCATCTCCAC 697  
 |||| |  
 CY 804 GAACAAGAATGTGGGGAITGTACAGCCCCGIGAGCATCTCTGGGACATCAAGCAGGGGCCCAA 863  
 |||| |  
 DB 698 CCNCGACGACGAGACGCTCACCAACGACGNATGCCGCCGACCTCAAGGACCGCTCAT 757  
 |||| |  
 CY 864 GGAGCCCTTCGGCGACTACGTGGACCGCTTCTTCAAGACCCT 905  
 |||| |  
 DB 758 CAAGCGGTCTATCCCCGACAACTACCTCGACGAGAAGCAT 799  
 |||| |

RESULT 5		851 bp	mRNA	linear	EST 08-APR-2003
CB645622					
DEFINITION	OSJNEB07H03.f OSJNEB Oryza sativa (japonica cultivar-group) cDNA clone OSJNEB7H03 5', mRNA sequence.				
ACCESSION	CB645622				
KEYWORDS	CB645622.1 GI:29640613				
SOURCE	EST.				
ORGANISM	Oryza sativa (japonica cultivar-group)				
	Oryza sativa (japonica cultivar-group)				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				
	Erihartoideae; Oryzaceae; Oryza.				
REFERENCE	1 (bases 1 to 851)				
AUTHORS	Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.				
TITLE	Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea				
JOURNAL	Unpublished				
COMMENT	Contact: Rod Wing Arizona Genomics Institute University of Arizona Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA Tel: 520 626 3967 Fax: 520 621 9288 Email: http://genome.arizona.edu				
	PCR Primers				
	FORWARD: gta aaa cga cgg cca gtg				
	BACKWARD: gga aac agc tat gat cat g				
	Plate: 07 row: H column: 03				
	Seq primer: gta aaa cga cgg cca gtg.				
FEATURES	Location/Qualifiers				
source	1..851				
	/organism="Oryza sativa (japonica cultivar-group)"				
	/mol_type="mRNA"				
	/cultivar="Nipponbare"				
	/db_xref="taxon:39947"				
	/clone="OSJNEB07H03"				
	/issue_type="Leaf"				
	/dev_stage="3 week"				
	/lab_host="DH10B"				
	/clone_lib="OSJNEB"				
	/note=vector: paluescript II KS +; Site_1: EcoRI; Site_2: XhoI; 24 hrs after inoculation with Rice Blast (Che 06061 )"				
BASE COUNT	170 a	325 c	238 g	118 t	
ORIGIN					
Query Match	5.3%; Score 78; DB 14; Length 851;				
Best Local Similarity	44.4%; Pred. No. 0.00074;				
Matches 312:	Conservative 0; Mismatches 350; Indels 0; Gaps 0;				
CY 204 CGACACCGGACCGGAGAGCTGAAGACCTGTTTCACACAGTGTGCCACATCCCTGTACTGGCT 263					
DB 105 GCAGATAGAAAATGGCGCGGAGAGCTTCCTCTTCACCTCCGAGTCGGTGAACGAGGGTCA 164					
CY 264 GCACGAGAAGATCGAGGTCCGCCACACCAAGGAGGCCCTGGACAAGATCGAGGAGGACCA 323					

Db 165 CCGGACAAGCTGTGCGACAGTGTGGACGCGGTGTGTGACGCGTGCCTGGGCCAGGA 224  
 QY 324 GAACAAGTGGCAGCAGAGATCCAGCAGCGCGAGCGCCCGCAGCAAGGCAAGGTGAGCCA 383  
 Db 225 CCGGACAGCAAGTGGGTGCGAGAGCTGCACCAAGACCAACATGTGTGTGTGTCGG 284  
 QY 384 GAATACGCCCTGTGAGAACCTGCGAGGCGAGCTGTGCACAGGCCATCAGCCCGCG 443  
 Db 285 CGAGATCACCACCAAGGCCACCGTCGACTAGGAGAAGATGTCGCGCAGACCTGCCCGCG 344  
 QY 444 CACCCTGACCCCTGGGTGAAGTGTGCGAGGAGAGGCGCTTACGCCCGCAGGTGAACCC 503  
 Db 345 CATCGGCTGTGTCGCGACAGCTGGCGCTCGAGCGCGACCGCTGCAAGGTGTGCTGCA 404  
 QY 564 CATGTTACCGCCCTGAGCGAGGCGCGACCCCGCCAGGACCTGAACAGATGTTGACAC 563  
 Db 405 CATCGACGACAGTCCGCGGACATCGCGAGCGGGTGCACGGCCATTCACCAAGCGCC 464  
 QY 564 CGTGGGGCGCCAGCGCGCATGCGATGCTGAAGGACACCATCAACGAGCAGCGCC 623  
 Db 465 CGAGGAGATCGCGCGCGCGACCGGCGCCACATGTCGGGTACGCCACCGAGAGACCC 524  
 QY 624 CGAGTGGGACCGCTGACCCCGCTGCGAGCGCGCGCCCATCGCCCGCGCAGATGCGCA 683  
 Db 525 CGAGCTGATGCGCTGACCGACAGCTGCTGCGCACCAAGCTCGGCGCGCGCTGCA 584  
 QY 684 GCGCGCGCGCAGCAGATGCGCGCGCACACAGCAGCTGCAGAGCAGATGCGCTGAT 743  
 Db 585 CCGCAAGAAGCGCACCTGCGCTGCTGAGCGCGCGAGCGGCAAGACCATCGCTGGA 644  
 QY 744 GACGAGAACCCCGCCTGCGCGCGCGGAGATCTAGAACGAGTGTGATGCTGCGCT 803  
 Db 645 GTAGCTCAACGACCGCGCGCGCATGTCGCGCTGCGCTGCGCGCGCTGCA 704  
 QY 804 GAACAAGATGCTGCGGATGTACAGCGCGCGTGTGAGGATCTGAGACATCAAGCAGCG 863  
 Db 705 CCACACAGCAGACGCTACCAACAGCAGAGATCGCGCGCGCGCTCAAGGAGCAG 764  
 QY 864 GGAGCGCTTGGGACTAGCTGGACCGCGCTTCTTCAAGACCT 905  
 Db 765 CAAGCGGTGATCCCGCAGAGTACTCTGAGCAGAGACCAT 806

RESULT 6  
 CB673355  
 LOCUS  
 DEFINITION OSJNE07N10.f OSJNE Oryza sativa (japonica cultivar-group) cDNA  
 clone OSJNE07N10.5, mRNA sequence.

ACCESSION CB673355  
 VERSION CB673355.1 GI:29677080

KEYWORDS  
 SOURCE  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 807)  
 Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
 Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
 Large-scale identification of ESTs involved in the interaction  
 between rice and Magnaporthe oryzae

JOURNAL Unpublished

COMMENT Contact: Rod Wing  
 Arizona Genomics Institute  
 University of Arizona  
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
 85721-0088, USA  
 Tel: 520 626 3967  
 Fax: 520 621 5288  
 Email: <http://genome.arizona.edu>  
 PCR Primers

FORWARD: gta aac cga cga cga gta  
 BACKWARD: gga aac agc tat gac cat g

Plate: 07 row: N column: 10  
 Seq primer: gta aac cga cga cga gta  
 Location/Qualifiers  
 1..807  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="OSJNE07N10"  
 /tissue\_type="Leaf"  
 /dev\_stage="3 week"  
 /lab\_host="DH10B"  
 /clone\_lib="OSJNE"  
 /note="Vector: pBluescript II KS +; Site1: EcoRI; Site2:  
 XhoI; 24 hrs after inoculation with Rice Blast (70-15)"  
 BASE COUNT 167 a 301 c 229 g 110 t  
 ORIGIN

Query Match 5.2%; Score 76.6; DB 14; Length 807;  
 Best Local Similarity 44.3%; Pred. No. 0.0013;  
 Matches 310; Conservative 0; Mismatches 389; Indels 0; Gaps 0;  
 QY 204 GCAGACCGGCGAGGAGCTGAAGAGCCTGTTCACACCGTGGCCACCCCTGATGCT 263  
 Db 105 GCAGATAGAAATGGCGCGAGAGAGCTTCCTCTTCACTCCGATCGTGAAGAGGGA 166  
 QY 254 GCAGGAGAGATCGAGTCCGCGACACAGAGGAGGCGCTGGACAGATCGAGCAGGA 323  
 Db 159 CCGGACACAGCTGTGGACACAGGTGTGGACCGCGGTCTGACGCGTGGCTCGCC 228  
 QY 324 GAACAAGTCCAGCAGAGATCCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 383  
 Db 229 CCGCGACAGCAAGTGGCTGCGAGAGCTGCACCAAGACCAACATGGTGTGTCGG 288  
 QY 384 GACACACCGCTGCTGACAGACCTGCGAGGCGAGATGTGTCACAGGAGCCATCAG 443  
 Db 289 CGAGATCAACACCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 348  
 QY 444 CACCTGGAACGCTGGTGAAGTGTATCGAGAGAGGCGCTTCAGCGCGCGAGTGC 503  
 Db 349 CATCGGCTTGTGTCGAGCAGCTGCGCGCTGACGCGCGCGCGCGCGCGCGCGCG 408  
 QY 504 CATGTTACCGCGCTGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 563  
 Db 405 CATCGAGCAGCAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 468  
 QY 564 CGTGGCG 623  
 Db 465 CGAGGAGATCG 528  
 QY 624 CGAGTGGGACCGGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 683  
 Db 529 CGAGCTGATGCGCGCTGAGCAGCTCTCGCGCGCGCGCGCGCGCGCGCGCGCG 588  
 QY 684 GCG 743  
 Db 589 CGCAACAACAGCGACCTGCGCTGCTAGCGCGCGCGCGCGCGCGCGCGCGCGCG 648  
 QY 744 GACCAACAACCG 803  
 Db 649 GTACCTCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 708  
 QY 804 GAACAAGATGCTGCGGATGTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 863  
 Db 709 CCAGCAGCAGGAGCGCTACCAACAGCAGAGATCGCGCGCGCGCGCGCGCGCG 768  
 QY 864 GGAGCGCTTCCGCGACTACGTGGAACCGCTTCTTCAAGAC 902  
 Db 769 CAAGCGGTGATCCCGCGACAAGTACCTCGCAGCAGAGAC 807

RESULT 7  
 CB681973

```

LOCUS      CB681973              812 bp      mRNA      linear      EST 09-APR-2003
DEFINITION OSJNEF08M09.f OSJNEF Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEF08M09 5', mRNA sequence.
ACCESSION  CB681973
VERSION     CB681973.1 GI:29685498
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartidae; Oryzaceae; Oryza.
REFERENCE   1 (bases 1 to 812)
AUTHORS     Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE       Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL     Unpublished
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 08 row: M column: 03
Seq primer: gta aaa cga cgg cca gtg.
            Location/Qualifiers
FEATURES             source
     source            1..812
                        /organism="Oryza sativa (japonica cultivar-group)"
                        /mol_type="mRNA"
                        /cultivar="Nipponbare"
                        /db_xref="taxon:39947"
                        /clone="OSJNEF08M09"
                        /tissue_type="Leaf"
                        /dev_stage="3 week"
                        /lab_host="DH10B"
                        /clone_lib="OSJNEF"
                        /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                        XhoI; Uninfected Control"
BASE COUNT          167 a 304 c 228 g 113 t
ORIGIN
Query Match          5.28; Score 76.4; DB 14; Length 812;
Best Local Similarity 44.3%; Pred. No. 0.0014;
Matches 311; Conservative C; Mismatches 391; Indels 0; Gaps 0;
QY 204 GCAGACCGGAGGAGGAGCTGAAGAGCCTGTTCAACACCGTGGCGGACCCCTGTACAGCT 263
Db 105 GCAGATAGAAATGGCGCGGAGACGTTCTCTTCACCTCGAGTCGCTGAAGAGAGGTCA 164
QY 264 GCACGAGAAGATCAGGTCCGGCACACCAAGAGGAGGCGCTGGACAAGATCGAGGAGAGCA 323
Db 165 CCGGACAAAGCTGTGCACACAGGTGTCCGACGCGGTGCTCGACGCTGCTGCCACAGTA 224
QY 324 GAACAAGTCCACAGCAAGATCCACAGCGCCGAGCGCGCGACGAAGGGCAAGGTGAGCA 383
Db 225 CCGGACAGCAAGGTGGCTGCGAGACGTGTCACCAAGACCAACATGGTGTGTTGG 284
QY 384 GAATACCCATGTCGACAACTGGAGCGCCAGTGGTGGACCCAGGCGCATCGGCGCCG 443
Db 285 CGAGATCACCACCAAGCGCCACCGCTCGACTACGAGAAGATCGTCCGCGGACACCTCGCGG 344
QY 444 CACCCCTGAAGCCCTGGGTGAAGTGTATCGAGGAGAAGGCGCTTCAGCCCGGAGGTGATCC 503
Db 345 CATGGCTTGGTGTCCGACCAAGCTGCGCTCGACGCGGACCGCTGCAAGTGTCTGTA 404
QY 504 CATGTTACCGCCTGAGCGGGCGGCGACCCCGCCAGGACCTGTAACAGATGTTGAAC 563
Db 405 CATGAGCAGCAGTCCGCCGACATCGCGGAGGGGTGGACCGGTCACCTTCACCAAGCGGCG 464

```

```

RESULT 8
CB684834
LOCUS      CB684834
DEFINITION OSJNEF14L24.f OSJNEF Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEF14L24 5', mRNA sequence.
ACCESSION  CB684834
VERSION     CB684834.1 GI:29688559
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartidae; Oryzaceae; Oryza.
REFERENCE   1 (bases 1 to 832)
AUTHORS     Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE       Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL     Unpublished
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 14 row: I column: 24
Seq primer: gta aaa cga cgg cca gtg.
            Location/Qualifiers
FEATURES             source
     source            1..832
                        /organism="Oryza sativa (japonica cultivar-group)"
                        /mol_type="mRNA"
                        /cultivar="Nipponbare"
                        /db_xref="taxon:39947"
                        /clone="OSJNEF14L24"
                        /tissue_type="Leaf"
                        /dev_stage="3 week"
                        /lab_host="DH10B"
                        /clone_lib="OSJNEF"
                        /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                        XhoI; Uninfected Control"
BASE COUNT          168 a 315 c 231 g 118 t
ORIGIN

```



```

Query Match          5.2%  Score 76.4;  DB 14;  Length 832;
Best local Similarity 44.1%;  Pred. No. 0.0014;
Matches 311;  Conservative 0;  Mismatches 391;  Indels 0;  Gaps 0;

QY 204 GCAGACCGGCGAGGAGGAGCTCAACAGAGCTGTCAACACCGGCGCCACCCCHGACTCGGT 263
DB 97 GCAGATGAATGGGCGGAGAGCTTCTCTTACCTCCGAGTCCGTGACGAGGGTCA 156
QY 264 GCACGAGAGATCGAGTTCGCGACACCAAGAGGCCCTGACAGATCGAGGAGGACA 323
DB 157 CCGGACCAAGCTGTCCGACAGGCTGCGAGCGGTGCTCGACGGTCCCTCCGCCAGGA 216
QY 324 GAACAAGTGGCAGCAGAGATCCACAGGCGGAGGCCGCCACAGGSCAAGTGAGGCA 383
DB 217 CCCCACAGCAAGTGGGCTGCGAGCTGACCAAGACCACTGTGTGATAGTGTTCGG 276
QY 384 CAATACCCCTATCGTSCAGAACTCTGACGGGCGAGATGTCACACGAGGCATCAGCGCGG 443
DB 277 CGAGATCACCAACAGGCGCACCTGAGTACGAGAAGATGTCGCGGACACTGCGCGCG 336
QY 444 CACCTGAACCGCTGGTGAAGGTGATCGAGGAGAGGCGCTTCAAGCCCGAGGTGATCCG 504
DB 337 CATCGCTTCGTGTCGAGAGAGCTGCGCTCGAGCGCCGACCGGTGCAAGGTGCTCTCAA 396
QY 504 CATGTTCACGCGCTGAGCGGGGCGGACCCCGCAGACCTGACACAGATGTTCAGAC 563
DB 397 CATGAGCAGAGTGCCTCCGATATGCGCAGGGGTGACCGGCCTTTCACCAAGCGGCC 456
QY 564 CTTGGGCGCCACCAAGCGGCATCGATGCTGAAGCAGACACCATCAAGGAGGAGCGCC 623
DB 457 CGAGAGATCGGCGCGCGACCGGCGACATGTTGGGCTACGCCACCGAGAGACCGC 515
QY 624 CGATGAGTGGCTGACAGCGCTGACAGCGCGGCGGCGGCTGCGCGCGCGCAGAGAGCGGA 683
DB 517 CGAGTGTATGCGCTGACAGCGCTGCTGCCACCAAGCTCGGCGCGCGCGCTCACCGAGT 576
QY 684 GCGCGCGCGCGGACATCGCGCGGACACCGACGCTGCGAGGAGGAGATGCGCTGCAT 743
DB 577 CGCAGAGAGGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 836
QY 744 GACGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 803
DB 637 GTACCTCAACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 896
QY 804 GAACAGATCTGCGGATGTACAGCGCGCTGAGCATCTGACATCAGAGGAGCGCGCGCA 863
DB 697 CGACGACGAGAGCGCTGACCAAGGAGGAGATGCGCGCGCGCGCGCGCGCGCGCGCGCG 756
QY 864 GGAGCGCTTCGCGGAGTACTGCTGAGCGCTTCTTCAAGACCGT 905
DB 757 CAAGCGGCTCATCTCCCGACAGTACTCTGACGAGAGACCAT 798

RESULT 9
BF259495
LOCUS
DEFINITION
BF259495
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

```

## JOURNAL COMMENT

for barley genomics: Morex unstressed seedling root cDNA library  
Unpublished  
On Nov 16, 2000 this sequence version replaced gi:11188608.  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Total hg bases = 388  
Seq primer: AATTAAACCTCAGTAAGGG  
High quality sequence stop: 758.  
location/qualifiers

## FEATURES source

1..759  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/cultivar="Morex"  
/db\_xref="taxon:112509"  
/clone="HVSMEf0019H19f"  
/tissue\_type="Seedling root"  
/lab\_host="TJ12"  
/clone\_lib="Hordeum vulgare seedling root EST library  
HVCNAC007 (Etisolated and unstressed)."  
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;  
Seeds were surface sterilized then germinated under axenic  
conditions in the dark at room temperature on filter paper  
with water, nystatin and cerotaxime in covered  
crystallization dishes. Five-day old seedling roots were  
then harvested, total RNA was prepared, poly(A) RNA was  
purified, one primary unamplified cDNA library was made,  
and 1 million pfu were in vivo excised to give pluescript  
SK(-) cDNA phagemids. These steps were performed in the  
close laboratory at the University of California,  
Riverside (Choi, Close, Fenton). Phagemids were plated and  
picked at the Clemson University Genomics Institute (CUGI)  
(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA  
preparations, DNA sequencing and sequence analysis were  
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates  
, Rambo, Main). The sequence has been trimmed to remove  
vector sequence and contains a minimum of 100 bases of  
phred value 20 or above. For more details on library  
preparation and sequence analysis see  
http://www.genome.clemson.edu/projects/barley. To order  
this clone see http://www.genome.clemson.edu/orders Also  
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)  
Genetically and physically anchored EST resources for  
barley genomics. Barley Genetics Newsletter 31:29-30.  
(http://wheat.pw.usda.gov/ggpages/bgw/31/cover.html)"

BASE COUNT 158 a 282 c 184 g 133 t 2 others

Query Match: 5.2% Score 76.2; DB 10; Length 759;  
Best local Similarity 47.7%; Pred. No. 0.0015;  
Matches 284; Conservative 0; Mismatches 305; Indels 6; Gaps 2;

QY 124 GAGAGTTCCTCGAACCCGCGCTGCGAGACCGAGCGGCTGCGAGCAGATCATC 183  
DB 13 GAGATACCCGCCCAACCCATCCATGGAACCTGATCAGCAGCTCGCGGACGAGGCCAGTCT 72  
QY 184 CGCAGCTGCACCCCGCCCTGCAGACCGCGGAGGAGCGTGTTCACACCC 243  
DB 73 CTCACATCGTCTCAAGCTCATCGGCGCCAAAGACCATGAGATCGCGCTCTACAG 132  
QY 244 GTGCCACCTCTACTGCGTGCACGAGAAGATCGAGTCCGCGACACCAAGAGGCGCTG 303  
DB 133 GGCTACTCCCTGCTCGCCACCGCGCTCGCATCCCGACGACGCCACCATCTTGGCCATG 192  
QY 304 GACAAGATCAGGAGGAGGAGAGACAGTGCACGACAGATCCAGAGGCCCGGCGCGCC 363  
DB 193 GACATCAACCGCGGAGAACTACGAGTGGGCTCCGCTGCATCGAAGGCGCGCGTGGCG 252  
QY 364 GACAAGGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAACCTGCAGGCGGCGAGTGGTG 423

Db	253	CACAAGATCGACTTCGCGGAGGCGCCGCGGCTC---CCGTCCTGGAGGCGCTTCGAG	309
QY	424	CACCAGGCCATCAGCCGCCGACCCCTGAACGCGTGGGTGAAGG-GATCGAGGAGAAAGGC	483
Db	310	GACGAGCCCAACCAACGCGACCTTCGACTTCGCTCTCG-GGACCGGACACGAGCAGC	369
QY	484	TCAGCCCGGAGGTGATCCCATGTTCACGGCCCTGAGCGAGGCGGCGACGCCCGCAGGAC	543
Db	370	CTCAACTACCAACGAGCGGCTCATGAACCTCTGCAAGGTCGCGGCGCTTCCTCGGCTAC	429
QY	544	CTGAACACGATGTGAACACCGGTGGCGGCGCACAGCGGCCATCAGATGCTGAAGAGAC	603
Db	430	--AACACCTCTGGAAGCGCTCCGCTGCTCTCCCGCGGAGCGGCCCA-GCGCAATAC	486
QY	504	ACCATCAACGAGGAGGCGCGGAGTGGGACGGGTGCAACCCCTGCACGGTGGCGGCATC	663
Db	487	ATCCGCTACTACGCGGACCTCGTTCCTCGACCTCAACAAGCGCTCGCGGAGCAGCGC	546
QY	664	GCCTCCGCGGACAGTGGCGGAGCTCGCGGACGAGATCGCGGACGACATCGCGACGAC	718
Db	547	GTNAGATCTGCCAGCTCCCGCTGGCGAGCGCATACCCCTCTGCCCGCGGCA	601
RESULT 10	BF484304		
LOCUS			
DEFINITION		562 bp mRNA linear EST C6-DEG-2000	
ACCESSION		WHE2321.F12_K23/2S Wheat pre-anthesis spike cDNA library Triticum	
VERSION		aestivum cDNA clone WHE2321.F12_K23, mRNA sequence.	
KEYWORDS		BF484304	
SOURCE		BF484304.1 GI:11567605	
ORGANISM		Triticum aestivum (bread wheat)	
REFERENCE			
AUTHORS		1 (bases 1 to 562)	
TITLE		Triticum aestivum (bread wheat)	
JOURNAL			
COMMENT			
FEATURES			
source			

```

Seq primer: 17.
Location/Qualifiers
1..649
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/strain="cv tokai"
/db_xref="GABI:549289"
/db_xref="taxon:112509"
/clone="HC15J04"
/tissue_type="Root"
/dev_stage="3 week old"
/clone_lib="HC"
/note="6 and 10 hour drought stress by placing plants on
moist paper (75% rel. humidity) in light"
BASE COUNT 128 a 253 c 175 g 92 t 1 others
ORIGIN

Query Match 5.1%; Score 75.6; DB 14; Length 649;
Best Local Similarity 47.7%; Pred. No. 0.0018;
Matches 284; Conservative 0; Mismatches 305; Indels 6; Gaps 2;

QY 124 GAGAAATTCGGCTTAACACCGCGCTGPGKAGACAGCGAGGCTGGAAGAGATATC 183
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
45 GAGATCACCGCAACACCACTGGAACTGATGACGAGCTGCGCGAGCGAGGCGCATTC 194
QY |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
184 GCGCAGCTGACCGCGCTGCGAGACCGCAGCGAGGAGCTGAAGAGCTCTGTGAACAC 243
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
105 CTCACATGCTCTTAAGCTCTATCGGCGCCAGAGACATGAGATCGGGCTTACACG 164
QY 244 GTGGCCACCTGTACTGCTGACGAGAGATCGAGTCCGCGACACCAAGGAGCGCTG 303
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
165 GGCTACTCCTGCTGCCACCGCGCTCGCATCCCGAGCGAGCGACCATCTTGGCATG 224
QY 304 GACAAATCGAGGAGGAGACAGACAGTCCAGCAGAGATCCACGCGCGAGCGCGCC 363
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
225 GACATCAACCGCAGAACTACAGTGGGGCTGGCGTGCATCGAAGCGCGCGCTGCGG 284
QY 364 GACAAGGGGAAGGTGAGGCGAGAACTACCCCATCGTGCAGAACTCGAGGCGCAGATGGTG 423
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
285 CACAAGTCTGCTTCAGCGAGCGCGCGGCTC---CGCTCTGGAGCGCTCTCTCAG 341
QY 424 CACAGGCTATAGCGCGCGCGACCTGAACGGCTGGTGAAGTGTATCGAGGAGAGGCT 483
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
342 GACGAGGCGCAACCAACGCGACCTTCGACTTCGCTTCTGTGACGCGGAGACAGCACTAC 401
QY 484 CTCAGCGCGAGGTATCGCGCTGTTCAAGCGCTTGAGCGAGGCGCGACGCGCGAGGAC 543
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
402 CTCAGTACACAGGCGCTCATGAAGTCTGTCAGGTCGGCGGCTCTCTGAGCTATGAC 461
QY 544 CTGAACAGATGTTGAACAACGTTGCGCGCGCGACACGAGCGCGCATGCAGATGCTGAAGGAC 603
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
462 ---AAGACCTCTGGAACGGCTCGCTGCTGCTCGCGCGAGCGCGCGCGCAAGTAC 518
QY 604 ACCATCAAGGAGGCGCGCGAGTGGACCGGCTGACCGGCTGACCGCGCGCGCGCATC 563
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
519 ATCGGTACTACCGGACTTCGTCCTCGACCTCAAGAGCGCGCTCGCGCGCGCGCGCG 578
QY 664 CGCGCGCGCGAGTGGCGGAGCGCGCGCGGAGGAGATCGCGGCGGACCGACCGCGCG 718
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
579 GTCGAGATCTGCCAGCTCCCGCTGCGGAGCGGATCACCTCTCTCGCGCGCGCGCA 633

RESULT 12
CB678700 834 bp mRNA linear EST 09-APR-2003
LOCUS OSJNEF01C19.f OSJNEF Oryza sativa (japonica cultivar-group) cDNA
DEFINITION clone OSJNEF01C19 5', mRNA sequence.
ACCESSION CB678700
VERSION CB678700.1 GI:29682425
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 834)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 21G088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gta
BACKWARD: gga aac agc tat gac cat g
Plate: 01 Row: C column: 19
Seq primer: gta aaa cga cgg cca gta.
Location/Qualifiers
1..834
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEF01C19"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEF"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Uninfected Control."
BASE COUNT 166 a 316 c 233 g 117 t 2 others
ORIGIN

Query Match 5.1%; Score 75.4; DB 14; Length 834;
Best Local Similarity 44.2%; Pred. No. 0.0021;
Matches 310; Conservative 0; Mismatches 392; Indels 0; Gaps 0;

QY 204 GCAGCGCGGAGGAGGAGCTGAAGAGCGCTGTCAACACCGTGTTCACACCGTGCACCTGTACTGGT 263
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
102 GCAGATAGAATCGCGGAGAGCGTTCCTCTTACCTCCGAGTCCGTTGAACGAGGTCA 161
QY 264 GCAGGAGAAGATCGAGTCGCGCACACCAAGCGCTGGACAACTGAGAGGAGGAGCA 323
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
162 CCGGAGCAAGCTGTGCGACAGGTGTGCGACGGTGTGCGACCGGTGCTCGACCGCTGCGCCAGGA 221
QY 324 GAACAAGTGCAGCAGAAAGATCCAGCAGGCGCGAGGCGCGGACAAAGGGAAGGTGAGCCA 383
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
222 CCGCGACAGCAAGTGGCTGCGAGAGCTGCACCAAGAGCAACATGGTGTGATGTTTCGG 281
QY 384 GAATACCCCATCGTCAGAACCTGAGGCGCGAGATGGTGCACAGCGCATGAGCCCGG 443
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
282 CGAGATCACCAAGCGCCACCTGCGACTACGAGAGATCGTCCGCGACCTGCGCGCG 341
QY 444 CACCTTGAACCGCTGGGTGAAGTGTATCGAGGAGAGAGGCTTCAGCCCGGAGGTGATGCC 503
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
342 CATCGGCTTCGTGTCGAGACGCTGCGCTCGAGCGCGGACCGCTGCAAGGTGCTGTCTCAA 401
QY 504 CATGTTCAACCGCTGTAGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 563
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
402 CATCGAGCAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 461
QY 564 CGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 623
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
462 CGAGGAGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 521
QY 624 CGAGTGGGAGCGGTGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 683
DB |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||
522 CGAGCTATCGCCCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 581

```

```

QY 684 GCCCGCGGAGACATCGCGGACACACAGCAGCAGCCTCGAGAGACAGATCCGCTGGAT 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 582 CCGAAGAACGGCAGCTGGCGCTGGCTGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAG 641
QY 744 GACGAGACACCCCGGATCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 803
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 642 GTACCTCAACAGCAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 701
QY 804 GAACAGATCGTGGGATGTCACAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 702 CCAGCAGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 761
QY 864 GGAGCGCTTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 905
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 762 CAGCGGCGATCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 803

RESULT 13
CB673814
LOCUS
DEFINITION
OSJNE08117.7 f. OSJNE Oryza sativa (japonica cultivar-group): cDNA
clone OSJNE08117 5', mRNA sequence.
ACCESSION
CB673814
VERSION
CB673814.1 GI:29677539
KEYWORDS
EST.
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 840)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
UNPUBLISHED
CONTACT: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR PRIMERS
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac aac tat gac cat g
Plate: 08 row: 1 column: 17
Seq primer: gta aaa cga cgg cca gtc.
FEATURES
Location/Qualifiers
1..840
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/db_xref="taxon:39947"
/clone="OSJNE08117"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNE"
/Note="Vector: phuescript II KS +; Site_1: EcoR; Site_2:
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"
BASE COUNT
169 a 318 c 233 g 119 t
ORIGIN
Query Match 5.1%; Score 75.4; DB 14; Length 840;
Best Local Similarity 44.2%; Pred. No. 0.0021;
Matches 310; Conservative 0; Mismatches 392; Indels 0; Gaps 0;
QY 204 GCAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 105 GCAGATAGAAATGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 164

```

```

QY 264 GCAGAGAGATCAGGTCCGCGACACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 165 CCGGAGCAAGTGTSCGACCAAGTGTGCGGAGCGGTGCTCGAGCGCTGCTCGGAGGAG 224
QY 324 GAACAAGTGCAGCAGAAAGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 225 CCGGAGCAGCAAGTGTGCGTGGAGAGCTGACCAAGACCAAGATGATGATGATGATG 284
QY 444 GAACTACCCCACTGAGAGACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 285 CGAGATCACCACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 344
QY 444 CACCTGNAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 345 CATCGGCTTGGTTCGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 404
QY 504 CATGTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 405 CATCGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 464
QY 564 CGTGGGCGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 445 CGAGGAGATCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 524
QY 624 CGAGTGGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 525 CGAGCTGATGCGGCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 584
QY 684 GCGCGGGGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 565 CCGGAAGAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 644
QY 744 GACGAGCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 803
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 645 GTACCTCAACGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 704
QY 804 GAACAAGATCGTGGGATGTCACAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 705 CCAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 764
QY 864 GGAGCGCTTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 905
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 765 CAGCGGCGATCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 806

RESULT 14
CB629976
LOCUS
DEFINITION
CB629976
VERSION
CB629976.1 GI:29624965
KEYWORDS
EST.
ORGANISM
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 766)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
UNPUBLISHED
CONTACT: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR PRIMERS

```

FORWARD: gta aaa cga cgg cca gta  
 BACKWARD: gga aac agc tat gac cat g  
 Plate: 06 ROW: K Column: 04  
 Seq primer: gta aaa cga cgg cca gta  
 Location/Qualifiers

## FEATURES

source  
 1..766  
 /organism="Oryza sativa (indica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="IR36"  
 /db\_xref="taxon:39946"  
 /clone="OSJNEC14E04"  
 /tissue\_type="Leaf"  
 /dev\_stage="3 week"  
 /lab\_host="DH10B"  
 /clone\_lib="OSLIB"  
 /note="Vector: pBluescript II KS +; Site:1: EcoRI; Site:2:  
 XhoI; 24 hrs after inoculation with Rice Blast (PO6-6-3)"  
 XhoI: 285 c 221 g 98 t

BASE COUNT 162 a 285 c 221 g 98 t

## ORIGIN

Query Match 5.1k; Score 74.8; DB 14; Length 766;  
 Best Local Similarity 44.2k; Pred. No. 0.0026;  
 Matches 310; Conservative 0; Mismatches 392; Indels 0; Gaps 0;  
 204 GCAGACGGCAGCAGGAGCTGAAGAGCTGTTCACACACCGTGGCCACCTGTACTGGG 263  
 56 GCAGATAGAAATGGCGGGAGACGCTTCCTCTCACCCTCCGAGTCCTGAACGAGGGTCA 115  
 264 GCAGAGAGATCGAGGCTGCGCAACCAAGAGCGCTTCACACAGATCGAGAGAGCA 323  
 116 CCGGACAAAGCTGTGCAACAGGCTTCGGACGCGTGTCTGACGCTGCTCGGACGGA 175  
 324 GAACAAGTCCAGCAGAAATCCAGCAGCGCCGAGCGCCGACAGGCAAGTGTAGGCA 383  
 176 CCGGACACAGGTTGGCTTCGAGNCGTGCACCAAGACCAACATGTGTATGGTTCGG 235  
 384 GAATACCCCATCTGTCGAGAACTCCAGGCGCAGATGTGTCACCAAGGCAATCAGCGCCG 443  
 236 CGAGATCACCAAGGCGCAGCTGACTACGAGAAGATCGTCCGCGACACTGCGCGCG 295  
 444 CACCTGAACGCTGGTGAAGATGATCGAGGAGAGCGCTTCAGCCCGGAGGTGATGCC 503  
 296 CATGGCTTGTGTCGAGAGGCTGGGCTTCAGCCGCGACGCTGCAAGGTGCTGCTAA 355  
 504 CATGTTACCGCCCTGAGCAGGGGCCACCGCCCGAGGACCTGAACAGATGTGAACAC 563  
 356 CATCGAGCAGCAGTCGCGCGACATCGCGCAGGGGTGCAGCGCCACTTCAACAGCGCC 415  
 564 CTTGGCGGCCACCAAGGCTGCTGATGATGCTGAAGACACCATCAGCAGGAGCGCG 623  
 416 CGAGGAGATCGCGCGCGGCGACAGGGGCAATGTTCCGCTACGCGACCGCAGAGCCG 475  
 624 CGAGTGGACCGCTGACACCGCTGACGCGCGCGCCCATCGCCCGCCAGATCGCGGA 683  
 476 CGAGCTGATGCCCTCAGCCAGCTCTCGCCACCAAGTTCGCGCGCCCTCAGCAGGT 535  
 684 GCGCGCGCGCAGCAGATGCGCGGACCAACACAGCAGCAGCTGAGGAGCAGATCGCTGAT 743  
 536 CCGCAAGAACGACCTGCGCGTGTCTAGCGCGCGCAGCGCAAGACCCAGGTCAACGTCGA 595  
 744 GACCAACACCGCCCATCGCGTGGGAGATCTACAGCGGTGGATCATCTCGGCT 803  
 596 GTACCTCAACAGCGCGCGCATGTTCCCGGTTCGCGTCCACACCGCTCTCTTCGAC 655  
 804 GAACAGATCTGCGGATGTACAGCGCGCTGAGCATCTCTGAGCATCAACAGGCGCCCAA 863  
 656 CCAGCAGGAGACCGCTCACCAACAGCAGAGATCGCCCGGACCTCAGAGGACGCTCAT 715  
 864 GGAGCCCTTCGCGACTAGCTGGACCGCTTCTTCAAGACCT 905  
 716 CAAGCGGGTCATCCCGGACAGTACCTCGACGAGAAGACCAT 757

RESULT 15  
 CB658304  
 LOCUS  
 DEFINITION

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

CB658304 797 bp mRNA linear EST 09-APR-2003  
 OSJNEC14E04.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA  
 clone OSJNEC14E04 5', mRNA sequence.  
 CB658304  
 CB658304.1 GI:29662029  
 EST.  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 1 (bases 1 to 797)  
 Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
 Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
 Large-scale identification of ESTs involved in the interaction  
 between rice and Magnaporthe grisea  
 Unpublished  
 Contact: Rod Wing  
 Arizona Genomics Institute  
 University of Arizona  
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
 85721-0088, USA  
 Tel: 520 626 3967  
 Fax: 520 621 9288  
 Email: http://genome.arizona.edu  
 PCR Primers  
 FORWARD: gta aaa cga cgg cca gta  
 BACKWARD: gga aac agc tat gac cat g  
 Plate: 14 ROW: K Column: 04  
 Seq primer: gta aaa cga cgg cca gta

## FEATURES

source  
 1..767  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="OSJNEC14E04"  
 /tissue\_type="Leaf"  
 /dev\_stage="3 week"  
 /lab\_host="DH10B"  
 /clone\_lib="OSJNEC"  
 /note="Vector: pBluescript II KS +; Site:1: EcoRI; Site:2:  
 XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"  
 XhoI: 300 c 225 g 110 t

BASE COUNT 162 a 300 c 225 g 110 t

## ORIGIN

Query Match 5.1k; Score 74.8; DB 14; Length 797;  
 Best Local Similarity 44.2k; Pred. No. 0.0026;  
 Matches 310; Conservative 0; Mismatches 392; Indels 0; Gaps 0;  
 204 GCAGCGCGCAGCAGGAGCTGAAGAGCTGTTCACACCGTGGCCACCTGTACTGGG 263  
 88 GCAGATAGAAATGGCGGGAGAGCTTCCTCTCACCCTCCGAGTCCGTGAACGAGGGTCA 147  
 264 GCAGAGAGATCAGATCGCGCAGCACACCAAGAGCGCTTCGACAGATCGAGAGAGCA 323  
 148 CCGGACAAAGCTGTGGACCAAGGTGTGGAGCGGTGTCTCGACGCGTGTCTCCCGCAGGA 267  
 324 GAACNAGTCCACGACAGATCAGCAGGCGCGGAGCGCGCGCAGCAAGGCAAGTGAACCA 383  
 208 CCGCGACAGCAAGGTGGCTGCGAGACGTGCAACCAAGACCAATGTGTGTGTTCGG 267  
 384 GAATACCCCATCTGTCAGAACCTGCAGGCGCGCAGATGGTGCACACAGGCGCATCAGCCCG 443  
 268 CGAGATCACCAAGCGCCACCTCTGACTACGAGAAGATGTCCTCGCAGACCTGCGCGCG 327  
 444 CACCTGAACCGCTGGGTGAAGGTGATCGAGGASAGGCTTCAGCGCGGAGGTGATGCC 503  
 328 CATCGGCTTGTGTCGAGCAGCGCTCGGCGCTCGAGCGCGGAGCGCTGCAAGTGTCTGTCAA 387  
 504 CATGTTACCGCGCTGAGCGAGGGCGCCACCCCGCAGGACCTGACACAGATGTTGAACAC 563

```

|||
388 CATCGAGCAGCAATCGCGGAGCATCGCGAGGCGGTGCACGGCCACTGCACCAAGCGCC 447
QY
564 CGTGGGGGGGACAGACGCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGAGG 623
|||
448 CGAGGAGATCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 507
QY
624 CGAGTGGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 683
Db
508 CGAGCTGATGCGCGCTCAGGACGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 567
QY
684 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 743
Db
568 CCGCAAGAACGGTACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 627
QY
744 GACGAGCAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 803
Db
628 GTACCTCAAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 687
QY
804 GACCAAGATCGTGGGATGTACAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 863
Db
588 CCAGCAGCAGGACCGTCAACCAACGACGAGATCGCGGCGGCGGCGGCGGCGGCGGCG 747
QY
864 GGACCCCTTCGGGAGTACGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 905
Db
748 CAAGCGGTGATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 789
```

Search completed: September 22, 2003, 22:50:10  
Job time : 3375.85 secs









```

ABL39953
ID ABL39953 standard; DNA: 1479 BP.
XX AC
XX AC
XX ABL39953;
XX
XX 15-MAY-2002 (first entry)
XX
XX Synthetic Gag polynucleotide sequence SEQ ID NO:3.
DE
XX
XX Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; net;
KW immunostimulant; gene therapy; gene; ds.
XX
XX Human immunodeficiency virus type C.
OS
XX Synthetic.
XX
XX W0200204493-A2.
XX
XX 17-JAN-2002.
XX
XX 05-JUL-2002: 2001W0-0521241.
XX
XX 05-JUL-2000: 2000US-0510313.
XX
XX (CHIR ) CHIRON CORP.
XX
XX (UYST-) UNIV STELLENBOSCH.
XX
XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
XX WPI: 2002-154920/20.
XX
XX
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful
XX in applications including DNA immunization or generation of packaging
XX cell lines, particularly in gene therapy
XX
XX Example 1: Fig 1: 233pp: English.
XX
XX The present invention describes expression cassettes comprising a
XX polynucleotide sequence encoding a polypeptide comprising immunogenic
XX HIV type C polypeptides. The expression cassettes comprise any of the
XX HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
XX Nef (I). (I) have immunostimulant activity and can be used in gene
XX therapy. The HIV type C polynucleotides are useful in applications
XX including DNA immunisation, generation of packaging cell lines, and
XX production of HIV type C proteins. The polynucleotides are particularly
XX useful in gene therapy and DNA immunisation applications. ABL39942 to
XX ABL40054 and ABB06204 to ABB06215 represent sequences used in the
XX exemplification of the present invention.
XX
XX Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 other;
XX
XX
XX Query Match 100.0%; Score 1479; DB 24; Length 1479;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-186;
XX Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ATGGGCGCGCGCCACATCTCGCGCGCGCGCAAGCTGGAGCGCTGGGAGCGCATCCG 60
XX
XX 1 ATGGGCGCGCGCGCCACATCTCGCGCGCGCGCAAGCTGGAGCGCTGGGAGCGCATCCG 60
XX
XX 61 CTGGCGCGCGCGCGCGCAAGAGTGTACATGATGAACACCTGTGTGTGGCGCGCGCGAG 120
XX
XX 61 CTGGCGCGCGCGCGCGCAAGAGTGTACATGATGAACACCTGTGTGTGGCGCGCGCGAG 120
XX
XX 121 CTGGAGAGTTCGCTTGACCCCGCGCTGCTGGAGACCGAGGCTGGAGAGAGATC 180
XX
XX 121 CTGGAGAGTTCGCTTGACCCCGCGCTGCTGGAGACCGAGGCTGGAGAGAGATC 180
XX
XX 181 ATCCGCGAGCTGACCCCGCGCTGCGAGACCGCGAGGAGCTGAAGAGCTGTTCAC 240
XX
XX 181 ATCCGCGAGCTGACCCCGCGCTGCGAGACCGCGAGGAGCTGAAGAGCTGTTCAC 240
XX
XX 241 ACCGTGGCCACCTGTACTCGCTGCGAGAGAGATCGAGTCCGCGACACCAAGAGGSC 300
XX
XX

```











QY	182	TCGCCAGCTGCACCCCGCCCTGCAGACCGGAGCGAGCTGACAGCTGTTCACCA	241
DB	1442	TCGCGACGTCACCCCGCCCTGCAGACCGGAGCGAGCTGACAGCTGTTCACCA	1501
QY	242	CGGTGGCCACCTGTACTCGTGCAGCAAGATCGAGTCCGGACACCAAGAGCGCC	302
DB	1502	CGGTGGCCACCTGTACTCGTGCAGCAAGATCGAGTCCGGACACCAAGAGCGCC	1561
QY	302	TGGCAAGATCGAGGAGCAGACAAGTGGCCAGCAAGATCCAGCAGCCGAGCCG	361
DB	1562	TGGCAAGATCGAGGAGCAGACAAGTGGCCAGCAAGATCCAGCAGCCGAGCCG	1621
QY	362	CGGACAAAGGCAAGTGCAGCAGAACTACCCATCGTGCAGCAAGTCCAGGCGCAGATG	421
DB	1622	CGGACAAAGGCAAGTGCAGCAGAACTACCCATCGTGCAGCAAGTCCAGGCGCAGATG	1681
QY	422	TGCACAGGCCATCAGCCCGCCGACCCCTGAAGGCCCTGGGTGAAGTGTATCGAGAGAAAG	481
DB	1682	TGCACAGGCCATCAGCCCGCCGACCCCTGAAGGCCCTGGGTGAAGTGTATCGAGAGAAAG	1741
QY	482	CGTTAGCCCGGAGGTGATCCCATGTTTCACAGCCCTGAGTCAGGCGGCACCCCGCAGG	541
DB	1742	CGTTAGCCCGGAGGTGATCCCATGTTTCACAGCCCTGAGTCAGGCGGCACCCCGCAGG	1801
QY	542	ACCTGAACACAGTGTGAACACCGTGGGGGGCCACCAGCCGCCCATGCAGATGCTGAAGG	601
DB	1802	ACCTGAACACAGTGTGAACACCGTGGGGGGCCACCAGCCGCCCATGCAGATGCTGAAGG	1861
QY	602	ACACCATCAAGCAGGAGGCGCGGANTGAGACCGGCTGTCACCCCTGCACGCCCGCTACA	661
DB	1862	ACACCATCAAGCAGGAGGCGCGGANTGAGAGCGCGTGCACCGCTGCACGCCCGCTACA	1921
QY	662	TGGCCCGCGGCAGATCGCGAGCCCGCGGCGCAGCCACACCGCCGCGCACACCGACCG	721
DB	1922	TGGCCCGCGGCAGATCGCGAGCCCGCGGCGCAGCCACACCGCCGCGCACACCGACCG	1981
QY	722	TGCAGGAGCAGATCGCCTGATGACACAGCAACCCCGCATCCCTCGTGGCGACATCTACA	781
DB	1982	TGCAGGAGCAGATCGCCTGATGACACAGCAACCCCGCATCCCTCGTGGCGACATCTACA	2041
QY	782	ACCGTGTGATCATCTGGGCTGAACAAGATCGTCGGATGTACAGCCCGTGGAGCATCC	841
DB	2042	ACCGTGTGATCATCTGGGCTGNACAAGAATCGTCGGATGTACAGCCCGTGGAGCATCC	2101
QY	842	TGGACATCAAGCAGGCGCCCAAGAGCCCTTCGCGGACTACCTGGACCGCTCTTTCAGA	901
DB	2102	TGGACATCAAGCAGGCGCCCAAGAGCCCTTCGCGGACTACCTGGACCGCGCTCTTTCAGA	2161
QY	902	CGCTCGCGCGCAGCAGACCCAGAGGTGAAGAACTGGATGACCCGACACCGTGGTGG	961
DB	2162	CGCTCGCGCGCAGCAGACCCAGAGGTGAAGAACTGGATGACCCGACACCGTGGTGG	2221
QY	962	TGCAGAAGCCCAACCCCGGACTGCAAGACCATCTCGCGCCTCTCGGCGCGGCGCCAGC	1021
DB	2222	TGCAGAAGCCCAACCCCGGACTGCAAGACCATCTCGCGCCTCTCGGCGCGGCGCCAGC	2281
QY	1022	TGGAGGAGATGATGACCGCCTGCCAGGCGCTGGCGCGCCCGCAGCCACAGCCCGCGTGC	1081
DB	2282	TGGAGGAGATGATGACCGCCTGCCAGGCGCTGGCGCGCCCGCAGCCACAGCCCGCGTGC	2341
QY	1082	TGGCGAGGCGATGAGCCAGCCCAACACCGCTGATGTCAGAGAGCAACTTTCAGG	1141
DB	2342	TGGCGAGGCGATGAGCCAGCCCAACACCGCTGATGTCAGAGAGCAACTTTCAGG	2401
QY	1142	GCCCGCGGCGATCGTCAAGTGCTTCAACTGCGGCGAAGGAGGCCACATTCGCGCGCAACT	1201
DB	2402	GCCCGCGGCGATCGTCAAGTGCTTCAACTGCGGCGAAGGAGGCCACATTCGCGCGCAACT	2461
QY	1202	GCGCGGCGCCCGCAAGAGGGCTGCTGGAAGTGGCGCAGGAGGCGCCACAGATGAGG	1261
DB	2462	GCGCGGCGCCCGCAAGAGGGCTGCTGGAAGTGGCGCAGGAGGCGCCACAGATGAGG	2521

QY	1262	ACTGCACGAGCGCCAGGCGCAACTTCTGTGGGCAAGATCTGTGGCCAGCCACCAAGGGCGGCC	1321
Db	2522	ACTGCACGAGCGCGGCGGCAACTTCTGTGGGCAAGATCTGTGGCCAGCCACCAAGGGCGGCC	2581
QY	1322	CGGCAACTTCTCTCAGAGCGCGCGGAGCGCCACCGCCCGCCCGCGGAGAGCTTCGGCT	1381
Db	2582	CGGCAACTTCTCTCAGAGCGCGCGGAGCGCCACCGCCCGCCCGCGGAGAGCTTCGGCT	2641
QY	1382	TCGAGGAGACCAACCCCGCGCCAGAGCAGAGCAGAGCAGGACCGCGAGACCTTGACCAGCC	1441
Db	2642	TCGAGGAGACCAACCCCGCGCCAGAGCAGAGCAGAGCAGGACCGCGAGACCTTGACCAGCC	2701
QY	1442	TGAAGAGCTGTTCGGCAACGAGCCCGCTGAGCCAGTAA	1479
Db	2702	TGAAGAGCTGTTCGGCAACGAGCCCGCTGAGCCAGTAA	2739
RESULT 9			
ACAQ3591	ID: ACAQ3591 standard; DNA; 5:84 BP.		
XX			
AC	ACAQ3591;		
XX			
DT	22-MAY-2003 (first entry)		
XX			
DE	Synthetic DNA encoding immunogenic HIV peptide #74.		
XX			
KW	Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;		
KK	vaccine; gene therapy; packaging cell line; humoral immune response;		
KK	cellular immune response; gene delivery vector; DNA immunisation;		
KK	ds.		
XX			
OS	Synthetic.		
XX			
PN	W02063004657-A1.		
XX			
FD	16-JAN-2003.		
XX			
PF	05-JUL-2002; 2002WO-US21421.		
XX			
PR	05-AUG-2001; 2001US-303192P.		
PR	31-AUG-2001; 2001US-316860P.		
PR	14-JAN-2002; 2002US-349728P.		
PR	16-JAN-2002; 2002US-34973P.		
PR	16-JAN-2002; 2002US-349871P.		
XX			
PA	(CHIR ) CHIRON CORP.		
XX			
PI	Zur Megede J, Barnett SW, Lian Y;		
XX			
DK	WPI; 2003-221602/21.		
XX			
PT	New synthetic polynucleotides encoding antigenic HIV type B and/or type		
PT	C polypeptides, useful as immunogenic compositions or vaccines for		
PT	generating humoral or cellular immune responses against HIV in a		
PT	subject, especially humans -		
XX			
ES	Example 1; Fig 79: 262pp; English.		
XX			
CC	The invention describes a synthetic polynucleotide encoding 2 or more		
CC	immunogenic HIV polypeptides, where at least 2 of the polypeptides are		
CC	derived from different HIV subtypes. The polynucleotide is useful for		
CC	immunisation, generation of packaging cell lines, or production of HIV		
CC	polypeptides. The polynucleotide and its encoded proteins are useful as		
CC	immunogenic compositions or vaccines for generating humoral or cellular		
CC	immune responses against HIV in a subject, or for inducing neutralising		
CC	antibodies against HIV. The gene delivery vector comprising the		
CC	polynucleotide is also useful for DNA immunisation of, or for		
CC	generating an immune response (e.g. a humoral or cellular immune		
CC	response) in, a subject such as a mammal, particularly a human. This		
CC	sequence encodes a human immunodeficiency virus immunogenic peptide.		
XX			
SQ	Sequence 5184 BP; 1139 A; 1852 C; 1610 G; 583 T; 0 other;		



Query Match 99.6% Score 1473.8; DB 25; Length 5184;  
Best Local Similarity 99.9%; Pred. No. 6, 2e-186;  
Matches 1475; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GGCGCGCGCGGCGAGATCTTGGCGCGCGGCAAGCTGGAGCGCTGGAGCGGATCCGGCT 62  
DB 1263 GGCGCGCGCGCGGAGATCTTGGCGCGCGGCAAGCTGGAGCGCTGGAGCGGATCCGGCT 1322  
QY 63 GGCGCGCGCGGCGAGATCTTACATGATGAAGCACTTGGTGGGCGAGCGGAGCT 122  
DB 1323 GGCGCGCGCGGCGAGATCTTACATGATGAAGCACTTGGTGGGCGAGCGGAGCT 1382  
QY 123 GGAGAGTTCCCTCTGAACCCCGGCTGCTGGAGACCGAGGAGCTGCAAGCAGATCAT 182  
DB 1383 GGAGAGTTCCCTCTGAACCCCGGCTGCTGGAGACCGAGGAGCTGCAAGCAGATCAT 1442  
QY 183 CCGCCAGTCCAGCCCGCTGCAGACCGGAGGAGCTGAAGAGCTGTCAACAC 242  
DB 1443 CCGCCAGTCCAGCCCGCTGCAGACCGGAGGAGCTGAAGAGCTGTCAACAC 1502  
QY 243 CGTGGCACCTCTACTGCTGCAGAGAGATCGAGTCCCGCACCAAGGAGCGCT 302  
DB 1503 CGTGGCACCTCTACTGCTGCAGAGAGATCGAGTCCCGCACCAAGGAGCGCT 1562  
QY 303 GGACAAGATCGAGGAGGAGCAACAAGTGGCAGCAGAGATCCAGGCGGAGCGCG 362  
DB 1563 GGACAAGATCGAGGAGGAGCAACAAGTGGCAGCAGAGATCCAGGCGGAGCGCG 1622  
QY 363 CGACAAGGCAAGTGGAGGAGCAACTACCCATCTGTGAGAGCTGCAGGCGGAGATGCT 422  
DB 1623 CGACAAGGCAAGTGGAGGAGCAACTACCCATCTGTGAGAGCTGCAGGCGGAGATGCT 1682  
QY 423 GCACAGGCGATCAGCCCGGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 482  
DB 1683 GCACAGGCGATCAGCCCGGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1742  
QY 483 CTTACGCGCGAGTGTATCCCATGTTTACCGCGCTTGAGCGGCGGCGGCGGAGGAG 542  
DB 1743 CTTACGCGCGAGTGTATCCCATGTTTACCGCGCTTGAGCGGCGGCGGCGGAGGAG 1802  
QY 543 CTTGACAGCATGTTGAACACCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 592  
DB 1803 CTTGACAGCATGTTGAACACCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1862  
QY 603 CACCATCAAGGAGGAGGCGGAGTGGAGCGGCTGAGCGGCGGCGGCGGCGGCGGCGGCGG 662  
DB 1863 CACCATCAAGGAGGAGGCGGAGTGGAGCGGCTGAGCGGCGGCGGCGGCGGCGGCGGCGG 1922  
QY 663 CGCGCGCGGCGGAGTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 722  
DB 1923 CGCGCGCGGCGGAGTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1982  
QY 723 GCAGAGCAGATGCGCTGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 782  
DB 1983 GCAGAGCAGATGCGCTGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2042  
QY 783 CGCGTGTATCATCTGGGCTGAAACAGATGCGGCGGATGTAAGCGGCGGCGGCGGCGGCGG 842  
DB 2043 CGCGTGTATCATCTGGGCTGAAACAGATGCGGCGGATGTAAGCGGCGGCGGCGGCGGCGG 2102  
QY 843 GGACATCAAGCAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 902  
DB 2103 GGACATCAAGCAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2162  
QY 903 CTTGCGCGCGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 962  
DB 2163 CTTGCGCGCGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2222  
QY 963 GCAGAGCGCAACCGGAGTGGAGAGCAGTCTGCGCGCTCTGCGCGCGGCGGCGGCGGCGG 1022  
DB 2223 GCAGAGCGCAACCGGAGTGGAGAGCAGTCTGCGCGCTCTGCGCGCGGCGGCGGCGGCGGCGG 2282

QY 1023 GGAGGAGATGATGACCGCTTGCAGGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1082  
DB 2283 GGAGGAGATGATGACCGCTTGCAGGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCT 2342  
QY 1083 GSCCGAGCGGATGAGCAGGCGCAACACCAAGCTGATGATGACAGAGAGCAAGCTTCAAGGG 1142  
DB 2343 GSCCGAGCGGATGAGCAGGCGCAACACCAAGCTGATGATGACAGAGAGCAAGCTTCAAGGG 2402  
QY 1143 CCCTCCGCGCATCGTCAAGTGTTCCTTCACTGGCGCAAGAGGAGGCGGCGGCGGCGGCGG 1202  
DB 2403 CCCTCCGCGCATCGTCAAGTGTTCCTTCACTGGCGCAAGAGGAGGCGGCGGCGGCGGCGG 2462  
QY 1203 CCCTCCGCGCGCGCAAGAGGCGCTGCTGAGAGTGGCGCAAGAGGAGGCGGCGGCGGCGG 1262  
DB 2463 CCCTCCGCGCGCGCAAGAGGCGCTGCTGAGAGTGGCGCAAGAGGAGGCGGCGGCGGCGG 2522  
QY 1263 CTGACCGGAGCGCGGCGCAAGTTCCTTGGCGCAAGATCTGGCGGCGGCGGCGGCGGCGG 1322  
DB 2523 CTGACCGGAGCGCGGCGCAAGTTCCTTGGCGCAAGATCTGGCGGCGGCGGCGGCGGCGG 2582  
QY 1323 CGGCAACTTCTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1382  
DB 2583 CGGCAACTTCTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2642  
QY 1383 CGAGGAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1442  
DB 2643 CGAGGAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1502  
QY 1443 GAAGAGCGCTTTCGCGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1479  
DB 2703 GAAGAGCGCTTTCGCGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2739

RESULT 10  
ACA03590  
T5 ACA03590 standard; DNA: 2742 BP.  
XX ACA03590;  
XX AC  
XX AT  
XX 22-MAY-2003 (first entry)  
XX  
XX Synthetic DNA encoding immunogenic HIV peptide #73.  
XX immunogenic HIV polypeptide; human immunodeficiency virus; HIV;  
KW vaccine; gene therapy; packaging cell line; humoral immune response;  
KW cellular immune response; gene delivery vector; DNA immunisation;  
KW ds.  
XX  
XX Synthetic.  
XX  
XX W0203004657-A1.  
XX  
XX 16-JAN-2003.  
XX  
XX 05-JUL-2002; 2002WO-US21421.  
XX  
XX 05-JUL-2001; 2001US-303192P.  
PR 31-AUG-2001; 2001US-316860P.  
PR 16-JAN-2002; 2002US-349728P.  
PR 16-JAN-2002; 2002US-349793P.  
PR 16-JAN-2002; 2002US-349871P.  
XX  
XX (CHIR ) CHIRON CORP.  
XX  
XX Zur Megede J, Barnett SW, Lian Y;  
PI  
XX  
XX WPI: 2003-221602/21.  
XX  
XX New synthetic polynucleotides encoding antigenic HIV type B and/or type  
PT C polypeptides, useful as immunogenic compositions or vaccines for  
PT generating humoral or cellular immune responses against HIV in a  
PT subject, especially humans -

Example 1; Fig 78; 262bp; English.

PS The invention describes a synthetic polynucleotide encoding 2 or more  
 XX immunogenic HIV polypeptides, where at least 2 of the polypeptides are  
 CC derived from different HIV subtypes. The polynucleotide is useful for  
 CC immunisation, generation of packaging cell lines, or production of HIV  
 CC polypeptides. The polynucleotide and its encoded proteins are useful as  
 CC immunogenic compositions or vaccines for generating humoral or cellular  
 CC immune responses against HIV in a subject, or for inducing neutralising  
 CC antibodies against HIV. The gene delivery vector comprising the  
 CC polynucleotide is also useful for DNA immunisation of, or for  
 CC generating an immune response (e.g. a humoral or cellular immune  
 CC response) in, a subject such as a mammal, particularly a human. This  
 CC sequence encodes a human immunodeficiency virus immunogenic peptide.  
 XX

SQ Sequence 2742 BP; 578 A; 1029 C; 859 G; 285 T; 0 other;

Query Match 59.6%; Score 1472.4; DH 25; Length 2742;

Best Local Similarity 99.9%; Pred. No. 1e-185;

Matches 1473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	2	TGGGCGCCGCGCCAGCATCTTCGCGGGGAGCTGGAGGCTGGAGGCTATCGCC	41
DB	1262	TCGGCGCGCGCCAGCATCTTCGCGGGGAGCTGGAGGCTGGAGGCTATCGCC	1321
QY	62	TGGCGCCCGCGCGCAAGAGTCTACATGATGAAGCAGCTGGTGGGCGCAGCGGAGC	121
DB	1322	TGGCGCCCGCGCGCAAGAGTCTACATGATGAAGCAGCTGGTGGGCGCAGCGGAGC	1381
QY	122	TGGAGAAGTTGGCCCTGAGACCCGCGCTGCTGGAGACAGGAGGCTGGAACAGATCA	181
DB	1382	TGGAGAAGTTGGCCCTGAGACCCGCGCTGCTGGAGACAGGAGGCTGGAACAGATCA	1441
QY	182	TCGCCAGCTGACCCCGCGCTTSCAGACCGCGAGGAGCTGAGAGCTGTTCACCA	241
DB	1442	TCGCCAGCTGACCCCGCGCTTSCAGACCGCGAGGAGCTGAGAGCTGTTCACCA	1501
QY	242	CGGTGGCGACCTGTACTGCTGCGAGAGAGATCGAGTCCGGACACACAGAGGCC	301
DB	1502	CGGTGGCGACCTGTACTGCTGCGAGAGAGATCGAGTCCGGACACACAGAGGCC	1561
QY	302	TGCAACAAGTTCGAGGAGAGCAGAACAGTGCACGACAGATCCAGCGCCGAGCGG	361
DB	1562	TGCAACAAGTTCGAGGAGAGCAGAACAGTGCACGACAGATCCAGCGCCGAGCGG	1621
QY	362	CGACAAAGGCGAGGTGAGCCAGACTACCCCATGCTGCAGAACTCGAGGCGCAGATGG	421
DB	1622	CGACAAAGGCGAGGTGAGCCAGACTACCCCATGCTGCAGAACTCGAGGCGCAGATGG	1681
QY	422	TGCACGAGGCGATCAGCCCGCCGACCCCTGAGCGCTGGTGAAGTGTATCGAGGAGG	481
DB	1682	TGCACGAGGCGATCAGCCCGCCGACCCCTGAGCGCTGGTGAAGTGTATCGAGGAGG	1741
QY	482	CTTTAGCCCGGAGGTGATCCCATGTTTACCGCGCTGAGCGAGGCGCCACCCCGCAGS	541
DB	1742	CTTTAGCCCGGAGGTGATCCCATGTTTACCGCGCTGAGCGAGGCGCCACCCCGCAGS	1801
QY	542	ACCTGAACACGATGTTGAACACCGTGGCGGCGCACAGCGCCCATGCAGATGCTGAAGS	601
DB	1802	ACCTGAACACGATGTTGAACACCGTGGCGGCGCACAGCGCCCATGCAGATGCTGAAGS	1861
QY	602	ACACCATCAACGAGAGGCGCGGAGTGGACCGGTGCACCCCGTGCACCGCGCGCCCA	661
DB	1862	ACACCATCAACGAGAGGCGCGGAGTGGACCGGTGCACCCCGTGCACCGCGCGCCCA	1921
QY	662	TCGCCCGCGCCAGATGCGCGAGCGCCGCGGAGCAGATCGCGGCGACACAGCACCC	721
DB	1922	TCGCCCGCGCCAGATGCGCGAGCGCCGCGGAGCAGATCGCGGCGACACAGCACCC	1981
QY	722	TGCAGGAGCAGATCGCCTGATGACGACCAACCCCGCATCCCGTGGGAGACATCA	781
DB	1982	TGCAGGAGCAGATCGCCTGATGACGACCAACCCCGCATCCCGTGGGAGACATCA	2041

QY	782	AGCGTGGATCATCTCGGCTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATCC	841
DB	2042	AGCGTGGATCATCTCGGCTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATCC	2101
QY	842	TGGACATCAAGCAGGCGCCCAAGAGCCCTTCGCGAC:ACGTGGACCGCTTCTTCAAGA	901
DB	2102	TGGACATCAAGCAGGCGCCCAAGAGCCCTTCGCGAC:ACGTGGACCGCTTCTTCAAGA	2161
QY	902	CCCTGCGCGCGAGCAGAGCAGCAGCAGAGGTGAAGAACTGGATGACGACACCGCTGCG	961
DB	2162	CCCTGCGCGCGAGCAGAGCAGCAGCAGAGGTGAAGAACTGGATGACGACACCGCTGCG	2221
QY	962	TGCAGAGCGCAACCCCGACTGCAAGACCATCTCGCGCTCTCGCGCCCGCGCGAGCC	1021
DB	2222	TGCAGAGCGCAACCCCGACTGCAAGACCATCTCGCGCTCTCGCGCCCGCGCGAGCC	2281
QY	1022	TGGAGGAGATGATGACCGCTGTCAGGCGCTGGCGCGCCAGCCACAGGCGCGCGTGC	1081
DB	2282	TGGAGGAGATGATGACCGCTGTCAGGCGCTGGCGCGCCAGCCACAGGCGCGCGTGC	2341
QY	1082	TGGCGGAGGCGATGAGCAGCGCCACACAGCGTGTATGTCAGAGAGCAACTTCAAGG	1141
DB	2342	TGGCGGAGGCGATGAGCAGCGCCACACAGCGTGTATGTCAGAGAGCAACTTCAAGG	2401
QY	1142	GCCCGCGCGCATGCTCAAGTGGCTTCAACTGCGGCAAGGAGGCGCCACATCGCGCGCACT	1201
DB	2402	GCCCGCGCGCATGCTCAAGTGGCTTCAACTGCGGCAAGGAGGCGCCACATCGCGCGCACT	2461
QY	1202	GCGCGCGCCCGCAAGAAGGGTGTCTGGAAGT:GCGGCAAGGAGGCGCCACAGATGAGG	1261
DB	2462	GCGCGCGCCCGCAAGAAGGGTGTCTGGAAGT:GCGGCAAGGAGGCGCCACAGATGAGG	2521
QY	1262	ACTGCACGAGCGCGCAGCGCAACTTCTTGGCAAGATCTGGCGCAGCCACCAAGGCGCGCC	1321
DB	2522	ACTGCACGAGCGCGCAGCGCAACTTCTTGGCAAGATCTGGCGCAGCCACCAAGGCGCGCC	2581
QY	1322	CGGCAACTTCTTGCAGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1381
DB	2582	CGGCAACTTCTTGCAGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	2641
QY	1382	TCGAGGAGCAGCCCG	1441
DB	2642	TCGAGGAGCAGCCCG	2701
QY	1442	TGAGAGGCTGTTCGCAACAGCCCTTGAGCCA	1475
DB	2702	TGAGAGGCTGTTCGCAACAGCCCTTGAGCCA	2735

RESULT 11

AA51625

15 AA51625 standard; DNA: 1479 BP.

XX AA51625;

XX AC

XX XX

DT 31-OCT-2000 (first entry)

XX HIV codon-optimized synthetic Gag polynucleotide.

DE Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;

XX DNA immunization; packaging cell line; antigen presentation; ss.

XX Human immunodeficiency virus type C strain AF110965.

OS Synthetic.

XX WO200039304-A2.

XX 36-JUL-2000.

XX 30-DEC-1999; 99WO-US31273.

XX 31-DEC-1998; 98US-0114495.

PR 01-SEP-1999; 99US-0152195.

XX PA (CHIR ) CHIRON CORP.  
 XX PI Barnett S, Zur Meede J;  
 XX DR WPI; 2000-452401/39.  
 XX PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV  
 XX PT Env polypeptide and the polypeptide useful for immunizing a mammal  
 XX PT especially human against HIV  
 XX PS Disclosure: Page 103-104; 113pp; English.  
 XX CC Expression cassettes comprising a polynucleotide encoding antigenic  
 CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are  
 CC useful in DNA immunization, generation of packaging cell lines and  
 CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag  
 CC expression cassettes exhibit increased potency for induction of  
 CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1  
 CC self-assemble into non-infectious virus-like particles which are used as  
 CC a matrix for the proper presentation of an antigen entrapped or  
 CC associated to the immune system of the host.  
 XX SQ Sequence 1479 BP; 325 A; 533 C; 461 G; 160 I; 0 other;

Query Match 98.5%; Score 1463; DB 21; Length 1479;  
 Best local Similarity 99.3%; Pred. No. 2e-184;  
 Matches 1469; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ATGGGCGCCGCGCCAGCATCTCGCGGCGGCAAGCTGGACGCTGGAGCGGTATCGGC 60  
 Db 1 ATGGGCGCCGCGCCAGCATCTCGCGGCGGCAAGCTGGACGCTGGAGCGGTATCGGC 60

Qy 61 CTGGGCGCGCGCGCGAAGAGTCTCATGATGATGATGATGATGATGATGATGATGATG 120  
 Db 61 CTGGGCGCGCGCGCGAAGAGTCTCATGATGATGATGATGATGATGATGATGATGATG 120

Qy 121 CTGGAGAGTTCGCTGAAACCGCGCGCTCTGGAGACGAGCGAGGCTGCAAGCAGAG 180  
 Db 121 CTGGAGAGTTCGCTGAAACCGCGCGCTCTGGAGACGAGCGAGGCTGCAAGCAGAG 180

Qy 181 ATGGGCGAGTCAGCCCGCGCTCGAGACGCGGAGCGGAGGCTGAAAGGCTGTTCAC 240  
 Db 181 ATGGGCGAGTCAGCCCGCGCTCGAGACGCGGAGCGGAGGCTGAAAGGCTGTTCAC 240

Qy 241 ACCGTGCGCCACCTGTACTGCTGTCAGAGATGATGATGATGATGATGATGATGATG 300  
 Db 241 ACCGTGCGCCACCTGTACTGCTGTCAGAGATGATGATGATGATGATGATGATGATG 300

Qy 301 CTGGACAGATCAGGAGGAGGAGACAGATGTCGACCAAGATGTCGAGGAGGAGGAG 360  
 Db 301 CTGGACAGATCAGGAGGAGGAGACAGATGTCGACCAAGATGTCGAGGAGGAGGAG 360

Qy 361 GCGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
 Db 361 GCGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420

Qy 421 GTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
 Db 421 GTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480

Qy 481 GCGTTCAGGCGCGGAGGTGATCCCATGTTTACCGGCTGAGGCGCTGAGGCGTGAAG 540  
 Db 481 GCGTTCAGGCGCGGAGGTGATCCCATGTTTACCGGCTGAGGCGCTGAGGCGTGAAG 540

Qy 541 GACCTGAACACGATGTTGAACACCGTGGGCGGACACGAGCGGCGGCGGCGGCGGCG 600  
 Db 541 GACCTGAACACGATGTTGAACACCGTGGGCGGACACGAGCGGCGGCGGCGGCGGCG 600

Qy 601 GACACCATCAACGAGGAGGCGGCGGAGTGGAGCGGCGGCGGCGGCGGCGGCGGCG 660  
 Db 601 GACACCATCAACGAGGAGGCGGCGGAGTGGAGCGGCGGCGGCGGCGGCGGCGGCG 660

Qy 661 ATCGCGCCGGCCAGATGCGGAGGCGCGCGGCGGAGCATCGCGGCGGAGCAGGAC 720  
 Db 661 ATCGCGCCGGCCAGATGCGGAGGCGCGCGGCGGAGCATCGCGGCGGAGCAGGAC 720

Qy 721 CTGAGGAGCAGATGCGCTGGATGACCAAGCCCGCCATCCCGTGGGCGACATTCAC 780  
 Db 721 CTGAGGAGCAGATGCGCTGGATGACCAAGCCCGCCATCCCGTGGGCGACATTCAC 780

Qy 781 AAGCGGTGGATCATCTGGGCGCTGAGCAAGATCGTGGCGCATGTACAGCCCGGTGAGCA 840  
 Db 781 AAGCGGTGGATCATCTGGGCGCTGAGCAAGATCGTGGCGCATGTACAGCCCGGTGAGCA 840

Qy 841 CTGACATCAAGCAGGCGCCCAAGGAGCCCTTCCGCGACTAGCTGAGCGCTTCTTCAAG 900  
 Db 841 CTGACATCAAGCAGGCGCCCAAGGAGCCCTTCCGCGACTAGCTGAGCGCTTCTTCAAG 900

Qy 901 ACCCTGCGCGCGGAGCAGACACCCAGGAGGCTTCCGCGACTAGCTGAGCGCTTCTTCAAG 960  
 Db 901 ACCCTGCGCGCGGAGCAGACACCCAGGAGGCTTCCGCGACTAGCTGAGCGCTTCTTCAAG 960

Qy 961 GTGCAAGAGCGCCAAACCCGAGTGCAGAGCATCTCTCGCGCTCTCGCGCGCGGCGGCG 1020  
 Db 961 GTGCAAGAGCGCCAAACCCGAGTGCAGAGCATCTCTCGCGCTCTCGCGCGCGGCGGCG 1020

Qy 1021 CTGAGGAGATGATGACCGCTTCCGAGGCGTGGGCGGCGGCGGCGGCGGCGGCGG 1080  
 Db 1021 CTGAGGAGATGATGACCGCTTCCGAGGCGTGGGCGGCGGCGGCGGCGGCGGCGG 1080

Qy 1081 CTGCGCGAGGCGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140  
 Db 1081 CTGCGCGAGGCGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140

Qy 1141 GCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200  
 Db 1141 GCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200

Qy 1201 TCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260  
 Db 1201 TCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260

Qy 1261 GATGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320  
 Db 1261 GATGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320

Qy 1321 CCGCGGAGCTTCTGTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380  
 Db 1321 CCGCGGAGCTTCTGTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380

Qy 1381 TTGAGGAGACACCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440  
 Db 1381 TTGAGGAGACACCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440

Qy 1441 CTGAAGAGCGCTGTCGCGCAAGCAGCCCTTGAAGCGAGTAA 1479  
 Db 1441 CTGAAGAGCGCTGTCGCGCAAGCAGCCCTTGAAGCGAGTAA 1479

RESULT 12  
 ABL39957  
 -D ABL39957 standard; DNA: 1479 BP.  
 XX AC ABL39957;  
 XX AC ABL39957;  
 XX DT 15-MAY-2002 (first entry)  
 XX Synthetic Gag polynucleotide sequence SEQ ID NO:20.  
 DE DE Human immunodeficiency virus type C; antigenic HIV type C protein;  
 KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;  
 KW immunostimulant; gene therapy; gene; ds.  
 XX OS Human immunodeficiency virus type C.  
 OS Synthetic.

XX WO20020493-A2.  
PN 17-JAN-2002.  
XX 05-JUL-2001; 2001WO-US21241.  
XX 05-JUL-2000; 2000US-0610313.  
XX (CHIR ) CHIRON CORP.  
PA (UYST-) UNIV STELLENBOSCH.  
XX Zur Mesede J. Barnett SW, Engelbrecht S, Van Rensburg EJ.  
XX WPI; 2002-154920/20.  
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful  
PT in applications including DNA immunization or generation of packaging  
PT cell lines, particularly in gene therapy  
XX Example 1; Fig 5; 23pp; English.  
XX The present invention describes expression cassettes comprising a  
CC polynucleotide sequence encoding a polypeptide comprising immunogenic  
CC HIV type C polypeptides. The expression cassettes comprise any of the  
CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or  
CC Nef (1). (1) have immunostimulant activity and can be used in gene  
CC therapy. The HIV type C polynucleotides are useful in applications  
CC including DNA immunisation, generation of packaging cell lines, and  
CC production of HIV type C proteins. The polynucleotides are particularly  
CC useful in gene therapy and DNA immunisation applications. ABL39942 to  
CC ABL40054 and ABL06204 to ABL05215 represent sequences used in the  
CC exemplification of the present invention.  
XX Sequence 1479 BP; 325 A; 533 C; 461 G; 160 T; 0 other;  
SQ

Query Match 98.9%; Score 1463; Db 24; Length 1479;  
Best Local Similarity 99.3%; Pred. No. 2e-184;  
Matches 1469; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGGGCGCGCGCGGAGCACTCTGGCGCGCGGAGCTGGAGCGCTGGAGCGCATCCGC 60  
DB 1 ATGGGCGCGCGCGGAGCACTCTGGCGCGCGGAGCTGGAGCGCTGGAGCGCATCCGC 60  
QY 61 CTGCGCGCGCGCGGAGCAAGTGTACATGATGATGAGCAACCTGGTGTGGCGGAGCGGAG 120  
DB 61 CTGCGCGCGCGCGGAGCAAGTGTACATGATGATGAGCAACCTGGTGTGGCGGAGCGGAG 120  
QY 121 CTGGAGAGTTTGGCCCTGAACCGCGGCTGTGGAGACCGAGGCTGCAAGCAGATC 180  
DB 121 CTGGAGAGTTTGGCCCTGAACCGCGGCTGTGGAGACCGAGGCTGCAAGCAGATC 180  
QY 181 ATCCGCCAGCTGCACCGCGCTCGAGACCGGAGCGAGGCTCAAGAGCTTTTCAAC 240  
DB 181 ATCCGCCAGCTGCACCGCGCTCGAGACCGGAGCGAGGCTCAAGAGCTTTTCAAC 240  
QY 241 ACCGTGGCGCAACCTTACTTGGTGCACAGAGATCGAGTCCGGACACCAAGAGGCC 300  
DB 241 ACCGTGGCGCAACCTTACTTGGTGCACAGAGATCGAGTCCGGACACCAAGAGGCC 300  
QY 301 CTGGACAGATCGAGGAGGAGCAGAACAAAGTCCACAGAGATCCACAGCGCGGAGGCC 360  
DB 301 CTGGACAGATCGAGGAGGAGCAGAACAAAGTCCACAGAGATCCACAGCGCGGAGGCC 360  
QY 361 GCGGACAGGGCAAGGTGAGCGAGAACTACCCCATCTGTGAGAGACCTGCGAGGCGCAATG 420  
DB 361 GCGGACAGGGCAAGGTGAGCGAGAACTACCCCATCTGTGAGAGACCTGCGAGGCGCAATG 420  
QY 421 GTGCACAGGCGCATCAGCGCGCGGACCTGAACCGCTGGGTGAAGGTGATTCAGAGGAG 480  
DB 421 GTGCACAGGCGCATCAGCGCGCGGACCTGAACCGCTGGGTGAAGGTGATTCAGAGGAG 480  
QY 481 GCCTTCAGCGCGGAGTGTATCCCATGTTTCCAGCGGCTGTGAGCGAGGCGCGCAACCGCGAG 540

DB 481 GCCTTCAGCGCGGAGTGTATCCCATGTTACCGCCCTGAGCGAGGCGCGCAACCGCGCAG 540  
QY 541 GACCTGAACAGCATGTTGAACACCGTGGCGCGCGCACCAAGCGCGCATGCAAGTGCCTGAAG 600  
DB 541 GACCTGAACAGCATGTTGAACACCGTGGCGCGCGCACCAAGCGCGCATGCAAGTGCCTGAAG 600  
QY 561 GACACCATCAAGAGAGGCGCGGAGTGGAGCGGCTGACACCGCTGACCGGCGCGCGCC 660  
DB 601 GACACCATCAAGAGAGGCGCGGAGTGGAGCGGCTGACACCGCTGACCGGCGCGCGCC 660  
QY 661 ATCGCGCGCGCGGAGATCGCGGAGCGCGGAGCGGAGCATCGCGCGCGCGCGCAGCACC 720  
DB 661 ATCGCGCGCGCGGAGATCGCGGAGCGCGGAGCGGAGCATCGCGCGCGCGCGCAGCACC 720  
QY 721 CTGCGAGGAGATCGCGCTGATGACGCAACCGCGCGCATCTCCCGTGGCGGAGCATCTAC 780  
DB 721 CTGCGAGGAGATCGCGCTGATGACGCAACCGCGCGCATCTCCCGTGGCGGAGCATCTAC 780  
QY 781 AAGCGGTGATCATCTCTGGCGCTGAACAAGATCGTGGCATGTACAGCGCGGTGAGCATC 840  
DB 781 AAGCGGTGATCATCTCTGGCGCTGAACAAGATCGTGGCATGTACAGCGCGGTGAGCATC 840  
QY 841 CTGGACATCAAGAGAGGCGCGCAAGAGCGCTTCCCGACTACGTGGACCGCTTTTCAAG 900  
DB 841 CTGGACATCAAGAGAGGCGCGCAAGAGCGCTTCCCGACTACGTGGACCGCTTTTCAAG 900  
QY 901 ACCCTGCGCGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 960  
DB 901 ACCCTGCGCGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 960  
QY 961 GTGCAAGAGCGCGCAACCGCGACTTCAAGACCATCTCTGCGCGCTCTGCGCGCGCGCGCAGC 1020  
DB 961 GTGCAAGAGCGCGCAACCGCGACTTCAAGACCATCTCTGCGCGCGCTCTGCGCGCGCGCAGC 1020  
QY 1021 CTGGAGGAGATGATGACCGCTGCGCAGGCGTGGCGCGCGCGCGCGCGCGCGCGG 1080  
DB 1021 CTGGAGGAGATGATGACCGCTGCGCAGGCGTGGCGCGCGCGCGCGCGCGCGCGG 1080  
QY 1081 CTGGCGGAGGCGATGAGCG 1140  
DB 1081 CTGGCGGAGGCGATGAGCG 1140  
QY 1141 GSCCG 1200  
DB 1141 GSCCG 1200  
QY 1201 TCG 1260  
DB 1201 TCG 1260  
QY 1261 GACTGCACCG 1320  
DB 1261 GACTGCACCG 1320  
QY 1321 GCG 1380  
DB 1321 GCG 1380  
QY 1381 TTGCGAGGAGCG 1440  
DB 1381 TTGCGAGGAGCG 1440  
QY 1441 CTGAGAGCGCTGTGCGCAACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1479  
DB 1441 CTGAGAGCGCTGTGCGCAACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1479

RESULT 23  
AAL44552  
ID AAL44552 standard: DNA; 1479 BP.  
XX  
AC AAL44552;

XX 08-NOV-2002 (first entry)

XX HIV-1 p55gag polypeptide coding sequence 3.

XX HIV; ds: vaccine; gene; immune response; microparticle;

XX adsorbent surface: poly(alpha-hydroxy acid); polyhydroxy butyric acid;

XX polycaprolactone; polyorthoester; polycyanoacrylate; detergent;

XX submicron emulsion; viral infection; bacterial infection;

XX parasitic infection; HIV-1 p55gag polypeptide.

OS Human immunodeficiency virus type 1.

XX Key Location/Qualifiers

PH variation replace (282, C)

FT /\*tag= a

FT variation replace (552, G)

FT /\*tag= b

FT variation replace (555, T)

FT /\*tag= c

FT variation replace (786, G)

FT /\*tag= d

FT variation replace (813, G)

FT /\*tag= e

FT variation replace (1002, T)

FT /\*tag= f

FT variation replace (1005, C)

FT /\*tag= g

FT variation replace (1092, G)

FT /\*tag= h

FT variation replace (1149, G)

FT /\*tag= i

FT variation replace (1158, C)

FT /\*tag= j

XX WC200226209-A2.

XX 04-APR-2002.

XX 28-SEP-2001; 2001WO-0530540.

XX 28-SEP-2001; 2000US-236105P.

XX 30-AUG-2001; 2001US-315905P.

XX (CHIR ) CHIRON CORP.

XX O'hagan D, Otten G, Donnelly CJ, Polo JM, Barnett S, Singh M;

XX Ulmer J, Dubensky TW;

XX WPI; 2002-519084/55.

XX A microparticle to which a biologically active macromolecule is

XX adsorbed, for use as a vaccine composition to treat viral, bacterial or

XX parasitic infections, comprises a polymer microparticle, a detergent

XX and a submicron emulsion.

XX Claim 72; Fig 5; 100pp; English.

XX The invention relates to a method of raising an immune response in a host

XX animal. The method of the invention comprises administering a

XX microparticle that has an adsorbent surface to which a first biologically

XX active macromolecule (e.g. a nucleic acid) has been adsorbed. The

XX microparticle comprises a polymer microparticle of poly(alpha-hydroxy

XX acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,

XX a polycyanoacrylate, a detergent, and submicron emulsion. The method/

XX microparticle of the invention is useful for immunising a host animal

XX against viral, bacterial or parasitic infections. The present DNA

XX sequence encodes a HIV-1 p55gag polypeptide.

XX Sequence 1479 BP: 325 A; 534 C; 460 G; 160 T; 0 other;

Query Match 98.8%; Score 1461.4; DB 24; Length 1479;

Best Local Similarity 99.3%; Pred. No. 3.2e-184;

Matches 1468; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATGGGCGCGGCGCCAGCATCTCTGGCGGGCGGAAGTGGACGCTGGGAGCCCATCCGC 60

DB : ATGGGCGCGGCGCGCATCTCTGGCGGGCGGAAGTGGACGCTGGGAGCCCATCCGC 60

QY 61 CTGGCGCCCGGCGCAAGAGTGTCTATCATATGAAGCACCCTGCTGTGGCCACCCGCGAG 120

DB 61 CTGGCGCCCGGCGCAAGAGTGTCTATCATATGAAGCACCCTGCTGTGGCCACCCGCGAG 120

QY 121 CTGGAGAAGTTGCGCTGAAACCCCGGCTCTGTGAGACACAGCGAGGCTGCAAGCAGATC 180

DB 121 CTGGAGAAGTTGCGCTGAAACCCCGGCTCTGTGAGACACAGCGAGGCTGCAAGCAGATC 180

QY 181 ATCGCGCAGCTGCACCCCGGCTCTGAGACCGGAGGAGAGCTGAAGAGCCTGTTCAAC 240

DB 181 ATCGCGCAGCTGCACCCCGGCTCTGAGACCGGAGGAGAGCTGAAGAGCCTGTTCAAC 240

QY 241 ACCGTGGCCACCTGTACTGCTGACGAGAAGATCGAGTCCGCGACACCAAGAGGAGCC 300

DB 241 ACCGTGGCCACCTGTACTGCTGACGAGAAGATCGAGTCCGCGACACCAAGAGGAGCC 300

QY 301 CTGACAAAGATCGAGAGGAGAGAGCAAGTGGCGAGAGATCCAGAGGCGCGAGGCC 360

DB 301 CTGACAAAGATCGAGAGGAGAGAGCAAGTGGCGAGAGATCCAGAGGCGCGAGGCC 360

QY 361 GCGACAAAGGCAAGGTGAGCGAGAACTACCCATCGTCGAGAACCTGAGGSCCAGATG 420

DB 361 GCGACAAAGGCAAGGTGAGCGAGAACTACCCATCGTCGAGAACCTGAGGSCCAGATG 420

QY 421 GTGACCAAGGCAATCAGCCCGGCGACCTTAAAGCCTGTGAAGGTGATCGAGGAGAG 480

DB 421 GTGACCAAGGCAATCAGCCCGGCGACCTTAAAGCCTGTGAAGGTGATCGAGGAGAG 480

QY 481 GCCTTCAGCCCGGAGGTGATCCCATCTTTCAGCGCCTGAGCGAGGCGGCCACCCCGCAG 540

DB 481 GCCTTCAGCCCGGAGGTGATCCCATCTTTCAGCGCCTGAGCGAGGCGGCCACCCCGCAG 540

QY 541 GACCTGAACACGATGTTGAACACCGTGGGGGGCGACCCAGCGCCCATGAGATGCTGAAG 600

DB 541 GACCTGAACACGATGTTGAACACCGTGGGGGGCGACCCAGCGCCCATGAGATGCTGAAG 600

QY 601 GACACCATCAACGAGGAGGCGCGGAGTGGGAGCGGTGCACCCGTCGACCGCGCGGCC 660

DB 601 GACACCATCAACGAGGAGGCGCGGAGTGGGAGCGGTGCACCCGTCGACCGCGCGGCC 660

QY 661 ATCGCCCGCGCGAGATGCGGAGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGCC 720

DB 661 ATCGCCCGCGCGAGATGCGGAGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGCC 720

QY 721 CTGAGGAGGAGATCGGCTGGAAGACAGAACCCCGGCGGAGCGGAGCGGAGCGGAGCGG 780

DB 721 CTGAGGAGGAGATCGGCTGGAAGACAGAACCCCGGCGGAGCGGAGCGGAGCGGAGCGG 780

QY 781 AAGCGGTGGATCATCTGGGCTGAACAAGATCTGGGAGTATACAGCCCGCTGAGCATC 840

DB 781 AAGCGGTGGATCATCTGGGCTGAACAAGATCTGGGAGTATACAGCCCGCTGAGCATC 840

QY 841 CTGGACATCAAGAGGCGCGGAGGAGCGGCTTCGCGACTACGTGGACCGCTTCTTCAAG 900

DB 841 CTGGACATCAAGAGGCGCGGAGGAGCGGCTTCGCGACTACGTGGACCGCTTCTTCAAG 900

QY 901 ACCGTGGCGCGGAGGAGAGAGCCAGGAGGTGAAGACTGGATGACCGACACCCCTGCTG 960

DB 901 ACCGTGGCGCGGAGGAGAGAGCCAGGAGGTGAAGACTGGATGACCGACACCCCTGCTG 960

QY 961 GTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020

DB 961 GTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020

QY 1021 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080

DB 1021 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080



```
QY 901 ACCCTGCGCGCGAGCAGACACCCAGAGAGTGAAGAACTGGATGACCGACACCCCTGCTG 960
Db 902 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
898 ACCCTGCGCGCGAGCAGACACCCAGAGAGTGAAGAACTGGATGACCGACACCCCTGCTG 957
QY 941 GTGACAGACCGCAGCCGCACTGCAAGACCATCTCTGCGCGCTCTCGCGCCGCGCGAGC 1025
Db 942 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
958 GTGACAGACCGCAGCCGCACTGCAAGACCATCTCTGCGCGCTCTCGCGCCGCGCGAGC 1017
QY 1021 CTGAGGAGATGATCAGCGCTCTGAGGCGCTGGCGCGCCAGCCAGCAAGCGCGCGCTG 1040
Db 1022 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
1018 CTGAGGAGATGATCAGCGCTCTGAGGCGCTGGCGCGCCAGCCAGCAAGCGCGCGCTG 1077
QY 1081 CTGCGCGAGCGATGAGCGAGCG---CAACACAGCGGATGATGACAAAGAGCAACTTC 1137
Db 1082 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
1078 CTGCGCGAGCGATGAGCGAGCGCAACACACAGCGCTGATGATCCAGAGCAACTTC 1137
QY 1138 AAGGCGCCCGCGCGATCGTCAAGTCTTCAACTCGCGCAAGGAGCGGCACATCGCGCGC 1197
Db 1139 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
1138 AAGGCGCCCGCGCGCGCTGGAAGTCTTCAACTCGCGCGCGAGGCGGCACATCGCGCGC 1197
QY 1198 AACTGCGCGCGCGCGCGAGAGCGCTCTGGAAGTGGCGCAAGGAGCGGCACAGATG 1257
Db 1199 AACTGCGCGCGCGCGCGAGAGCGCTCTGGAAGTGGCGCAAGGAGCGGCACAGATG 1257
QY 1258 AAGGACTGACGAGCGCGAGCGCAACTTCTGCGCAAGATCTGCGCAGCGCACAAGCGAC 1327
Db 1259 AAGGACTGACGAGCGCGAGCGCAACTTCTGCGCAAGATCTGCGCAGCGCACAAGCGAC 1327
QY 1318 CGCCCGCGCACTTCTGTCAGAGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGAGCTTC 1377
Db 1319 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
1318 CGCCCGCGCACTTCTGTCAGAGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGAGCTTC 1377
QY 1378 CGCTTCTGAGG-----AGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1425
Db 1379 CGCTTCTGAGG-----AGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1425
QY 1426 GAGACCGCTGACGAGCTGAAGAGCTGTGCGCAAGCAGCCCGCTGAGCGCACTAA 1475
Db 1438 GAGCGCGCTGACGAGCTGAAGAGCTGTGCGCAAGCAGCCCGCTGAGCGCACTAA 1491

RESULT 15
ABL39972
ID ABL39972 standard: DNA: 1494 BP.
XX
AC ABL39972:
XX
DT 15-MAY-2002 (first entry)
XX
DE Synthetic Gag polynucleotide sequence SEQ ID NO:51.
XX
KW Human immunodeficiency virus type C: antigenic HIV type C protein:
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX
OS Human immunodeficiency virus type C.
OS Synthetic.
XX
PN W020204493-A2.
XX
PD 17-JAN-2002.
XX
PF 05-JUL-2001: 2001WO-US21241.
XX
PR 05-JUL-2000: 2000US-0610313.
XX
PA (CHIR ) CHIRON CORP.
PA (OYST-) UNIV STELLENBOSCH.
XX
PI Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX WPI; 2002-154920/20.
XX
```

```
PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful
in applications including DNA immunization or generation of packaging
cell lines, particularly in gene therapy
XX
PS Claim 7: Fig 22: 233pp: English.
XX
CC The present invention describes expression cassettes comprising a
polynucleotide sequence encoding a polypeptide comprising immunogenic
HIV type C polypeptides. The expression cassettes comprise any of the
HIV type C sequences encoding gag, pol, vif, vpr, tat, rev, vpu, env or
nef (1). (1) have immunostimulant activity and can be used in gene
therapy. The HIV type C polynucleotides are useful in applications
including DNA immunisation, generation of packaging cell lines, and
production of HIV type C proteins. The polynucleotides are particularly
useful in gene therapy and DNA immunisation applications. ABL39942 to
ABL40054 and ABL406204 to ABL406215 represent sequences used in the
exemplification of the present invention.
XX
SQ Sequence 1494 BP: 125 A; 557 C; 455 G; 157 T; 0 other;
```

```
Query Match 87.1%; Score 1288; DB 24; Length 1494;
Best Local Similarity 92.9%; Prod. No. 2e-161;
Matches 1393; Conservative 0; Mismatches 80; Indels 27; Gaps 3;
QY 1 ATGGCGCGCGCGCGCGAGATCTCTGCGCGCGCGCAAGCTGGAGCGCGCGCGCGCGCG 60
Db 1 ATGGCGCGCGCGCGCGAGATCTCTGAGCGCGCGCAAGCTGGAGCGCGCGCGCGCGCG 60
QY 61 CTGCGCGCGCGCGCGCGAGAGTCTTACATGATGAGACCTGCTGTGGCGCGCGCGCGAG 120
Db 61 CTGCGCGCGCGCGCGCGAGAGTCTTACATGATGAGACCTGCTGTGGCGCGCGCGCGAG 120
QY 121 CTGACAGAGTCTGCGCTGAACCGCGCGCTCTGAGACCGAGCGAGGCTCAAGACGATC 180
Db 121 CTGACAGGCTTTCGCTGAACCGCGCTCTGAGACCGAGCGAGGCTCAAGACGATC 180
QY 181 ATCCCGCAGCTGACCGCGCGCTGACAGCGCGCGCGAGCGAGCGTGAAGAGCGCTTCAAC 240
Db 181 ATCAAGCAGCTGACCGCGCGCTGACAGCGCGCGCGAGCGAGGCTGCGCGCGCTTCAAC 240
QY 241 ACCGTGCGCGCGCTGCTACTGCTGACGAGAGATCGAGGTCGCGCGACACCAAGAGCGCC 300
Db 241 ACCGTGCGCGCGCTGCTACTGCTGACGAGAGATCGAGGTCGCGCGACCAAGAGAGCGCC 300
QY 301 CTGACAGAGTCTGAGGAGGAGAGAGTGGCCAGAGAGATCCAGAGCGCGCGAGCGCC 360
Db 301 CTGACAGAGTCTGAGGAGGAGAGAGTGGCCAGAGAGATCCAGAGCGCGCGAGCGCC 360
QY 361 GCGACAGAGGAGGAGGAGGAGTACCGCACTACCGCACTGCGAGAACCTGCGAGCGCGAG 420
Db 361 GCGACAGGAGGAGGAGGAGTACCGCACTACCGCACTGCGAGAACCTGCGAGCGCGAG 417
QY 421 GTGACCGAGGCGATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 418 GTGACCGAGGCGATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477
QY 481 GCGTTCAGCGCGCGAGGAGTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 478 GCGTTCAGCGCGCGAGGAGTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
QY 541 GACCTGACAGAGGAGTGTGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 538 GACCTGACAGAGGAGTGTGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
QY 601 GACACCATCAAGGAGGAGGCGCGCGAGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 598 GACACCATCAAGGAGGAGGCGCGCGAGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 657
QY 661 ATCCCGCGCGCGCGCGAGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Db 658 GTGCGCGCGCGCGCGAGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717
QY 721 CTGACGAGGAGAGATCGCGCTGGATGACCGAGCAACCGCGCGCGCGCGCGCGCGCGCG 780
```

```
Db      |||||||
718  CTGCAGGACGATGCCCTGGATGACCACACCCGCCATCCCGTGGAGACATCTAC 777
QY      |||||||
781  AAGCGGTGGATCATCTCGGCGTGAACAAGATCGTGGGATGTACAGCCCGGTGAGCATC 840
Db      |||||||
778  AAGCGGTGGATCATCTCGGCGTGAACAAGATCGTGGGATGTACAGCCCGGTGAGCATC 837
QY      |||||||
841  CTGGACATCAAGCAGGCGCCCAAGGAGCCCTCCCGGACTAGCTGGACCGCTTCTCAAG 900
Db      |||||||
838  CTGGACATCAAGCAGGCGCCCAAGGAGCCCTCCCGGACTAGCTGGACCGCTTCTCAAG 897
QY      |||||||
901  ACCTCGCGCGCGGAGCAGCAGCCAGGAGGTGAAGAACTGGATGACCGGACACCTGCTG 960
Db      |||||||
898  ACCTCGCGCGCGGAGCAGCAGCCAGGAGGTGAAGAACTGGATGACCGGACACCTGCTG 957
QY      |||||||
961  GTCCAGAACCCAAACCCCACTGCAAGACCATCTCTGGGCTCTGGCCCGCGGCGCAGC 1020
Db      |||||||
958  GTCCAGAACCCAAACCCCACTGCAAGACCATCTCTGGGCTCTGGCCCGCGGCGCAGC 1017
QY      |||||||
1021 CTGGAGGAGATGATACCGCTCTGCCAGGCGGTGGGCGGCCCGCAGCCACAAAGCCCGGCTG 1080
Db      |||||||
1018 CTGGAGGAGATGATACCGCTCTGCCAGGCGGTGGGCGGCCCGCAGCCACAAAGCCCGGCTG 1077
QY      |||||||
1081 CTGGCGGAGCGGATGAGCGAGGCCAACCCAGCGGTGATGATCAGAAAGAGCAACTTCAAG 1140
Db      |||||||
1078 CTGGCGGAGCGGATGAGCGAGGCCAACCCAGCGGTGATGATCAGAAAGAGCAACTTCAAG 1137
QY      |||||||
1141 GGCCCGCGGCGCATCGTCAAGTGCCTCAACTGGGCAAGGAGGACACATCGCCCGCAG 1200
Db      |||||||
1138 GGAGCAACCGGATCATCAAGTGCCTCAACTGGGCAAGGAGGACACATCGCCCGCAG 1197
QY      |||||||
1201 TCGCGGCGCCCGCGCAAGAGGCTCTGGAAGTGGGCAAGGAGGCGGACACAGATGAAG 1260
Db      |||||||
1198 TCGCGGCGCCCGCGCAAGAGGCTCTGGAAGTGGGCAAGGAGGCGGACACAGATGAAG 1257
QY      |||||||
1261 GACTGACCGAGCGCCAGGCGCAACTTCCTGGGCAAGATCTGGCCCGAGCCACAGGCGCG 1320
Db      |||||||
1258 GACTGACCGAGCGCCAGGCGCAACTTCCTGGGCAAGATCTGGCCCGAGCCACAGGCGCG 1317
QY      |||||||
1321 CCGGCAACTTCCTGGAGAGCGCGC-----CCGAGGCGCAAGCGG 1359
Db      |||||||
1318 CCGGCAACTTCCTGGAGAGCGCGC-----CCGAGGCGCAAGCGG 1377
QY      |||||||
1360 CCGCGCGCGGAGAGCTTCGCGCTTCGAGGAGACACCCCGGCGCAGAGCAGAGAGCAAG 1419
Db      |||||||
1378 CCGCGCGCGGAGAGCTTCGCGCTTCGAGGAGACACCCCGGCGCAGAGGAG---AAG 1434
QY      |||||||
1420 GAGCGGAGAGCTGACGAGCCCTGAAGAGCCCTTGGCCAAAGACCCCTGAGCCAGTAA 1479
Db      |||||||
1435 GAGCGGAGCGCCCTGACGAGCCCTGAAGAGCCCTTGGGAGGAGCCCTGAGCCAGTAA 1494
```

Search completed: September 22, 2003, 17:47:54

Job time : 432.683 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 17:26:03 ; Search time 5594.26 Seconds  
(without alignments)  
10515.605 Million cell updates/sec

Title: US-09-475-704A-3

Perfect score: 1479

Sequence: 1 atgagcgccgcgcagcat.....acgacccccgagccagtaa 1479

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs., 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*

1: gb-ba:\*

2: gb-htg:\*

3: gb-in:\*

4: gb-om:\*

5: gb-ov:\*

6: gb-pat:\*

7: gb-ph:\*

8: gb-pl:\*

9: gb-pr:\*

10: gb-ro:\*

11: gb-sts:\*

12: gb-sy:\*

13: gb-un:\*

14: gb-vi:\*

15: cm-ba:\*

16: cm-fun:\*

17: cm-hum:\*

18: cm-in:\*

19: cm-mu:\*

20: cm-or:\*

21: cm-ot:\*

22: cm-ov:\*

23: cm-pat:\*

24: cm-ph:\*

25: cm-pi:\*

26: cm-ro:\*

27: cm-sts:\*

28: cm-un:\*

29: cm-vi:\*

30: em-htg-hum:\*

31: em-htg-inv:\*

32: em-htg-other:\*

33: em-htg-mus:\*

34: em-htg-pin:\*

35: em-htg-rod:\*

36: em-htg-mam:\*

37: em-htg-vrt:\*

38: em-sy:\*

39: em-htgo-hum:\*

40: em-htgo-mus:\*

41: em-htgo-other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1479	100.0	1479	6	AX455887	AX455887 Sequence
2	1479	100.0	1479	6	AX468543	AX468543 Sequence
3	1473.4	99.6	1479	6	AX468547	AX468547 Sequence
4	1463	98.9	1479	6	AX455904	AX455904 Sequence
5	1268.8	87.7	1491	6	AX455963	AX455963 Sequence
6	1288	87.1	1494	6	AX455935	AX455935 Sequence
7	1276.8	86.3	1509	6	AX455888	AX455888 Sequence
8	1275.2	86.2	1509	6	AX468544	AX468544 Sequence
9	1273.6	86.1	1509	6	AX468548	AX468548 Sequence
10	1260.8	85.2	1509	6	AX455905	AX455905 Sequence
11	1221.2	82.6	9166	6	AX427930	AX427930 Sequence
12	1206	81.5	9788	6	AX427936	AX427936 Sequence
13	1205.4	81.5	4288	6	AX149648	AX149648 Sequence
14	1197.8	81.0	1509	12	AF201927	AF201927 Synthetic
15	1159.6	78.4	1548	12	AF287354	AF287354 Synthetic
16	1134.4	76.7	1539	6	AX306431	AX306431 Sequence
17	1131.8	76.5	6438	6	AX427919	AX427919 Sequence
18	1131.8	76.5	9194	6	AX427925	AX427925 Sequence
19	1131.8	76.5	9194	6	AX427926	AX427926 Sequence
20	1131.8	76.5	12411	6	AX427927	AX427927 Sequence
21	1120.4	75.8	9169	6	AX427931	AX427931 Sequence
22	1116.6	75.5	9170	6	AX427928	AX427928 Sequence
23	1116.6	75.5	9189	6	AX427921	AX427921 Sequence
24	1116.6	75.5	9792	6	AX427932	AX427932 Sequence
25	1114.6	75.4	1847	12	AF202464	AF202464 Synthetic
26	1114.6	75.4	1847	12	AF202465	AF202465 Synthetic
27	1110.6	75.1	9167	6	AX427933	AX427933 Sequence
28	1110.6	75.1	9782	6	AX427935	AX427935 Sequence
29	1110.6	75.1	9783	6	AX427934	AX427934 Sequence
30	1105.2	74.7	9785	6	AX427938	AX427938 Sequence
31	1091.8	73.8	4352	12	AF287352	AF287352 Synthetic
32	1087.2	73.5	4343	6	AX306428	AX306428 Sequence
33	1076.4	72.8	4341	6	AX306429	AX306429 Sequence
34	1074.8	72.7	3981	6	AX306430	AX306430 Sequence
35	1070.8	72.4	9407	6	AX427937	AX427937 Sequence
36	1066	72.1	4353	12	AF287353	AF287353 Synthetic
37	990.8	67.0	1905	6	AX457086	AX457086 Sequence
38	976.6	66.0	1532	6	BD057031	BD057031 Synthetic
39	949.4	64.2	4307	6	AX035453	AX035453 Sequence
40	949.4	64.2	4307	6	AX056836	AX056836 Sequence
41	947.6	64.1	4307	6	AR302563	AR302563 Sequence
42	947.6	64.1	4307	6	AX019132	AX019132 Sequence
43	947.6	64.1	4307	6	BD130468	BD130468 Antiviral
44	939.8	63.5	4307	6	AX283597	AX283597 Sequence
45	939.8	63.5	4307	6	AX283607	AX283607 Sequence

ALIGNMENTS

RESULT: 1  
AX455887  
LOCUS AX455887 Sequence 3 from Patent WO0204493, 1479 bp DNA linear PAT 06-JUL-2002  
DEFINITION Sequence 3 from Patent WO0204493.  
ACCESSION AX455887  
VERSION AX455887.1 GI:21714880  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1  
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.  
TITLE polynucleotides encoding antigenic hiv type c polypeptides.  
JOURNAL polypeptides and uses thereof  
Patent: WO 0204493-A 3 17-JAN-2002;

[illegible]

QY

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

Qy 1 ATGGGGCCCGCCAGCATCCTGCGCGCGCAAGCTGGACGCCCTGGAGCGCATCCGC 60

```
Db 1 ATGGGCGCCGCGCCAGCACTCTGCGGGGGGAAAGCTGGACGCTGGGAGGCAATCCGC 60
Qy 61 CTGGCCCGCGGCGCAAGAGTGTACATGATGAAGCACTGTGTGGGCGACGCGGAG 120
Db 61 CTGGCCCGCGGCGCAAGAGTGTACATGATGAAGCACTGTGTGGGCGACGCGGAG 120
Qy 121 CTGGAGAAGTTGCGCCCTGAACCCCGGCTGCTGGAGACCAAGGAGGCTGCAAGCAATC 180
Db 121 CTGGAGAAGTTGCGCCCTGAACCCCGGCTGCTGGAGACCAAGGAGGCTGCAAGCAATC 180
Qy 181 ATCGCCAGCTGCAACCCGCTGTCAGACCGGAGCGAGAGCTGAAGAGCTGTTCAC 240
Db 181 ATCGCCAGCTGCAACCCGCTGTCAGACCGGAGCGAGAGCTGAAGAGCTGTTCAC 240
Qy 241 ACCGTGGCCACCTGTACTGCTGCACGAGAAGATCGAGGTCCCGACACCAAGAGAGGC 300
Db 241 ACCGTGGCCACCTGTACTGCTGCACGAGAAGATCGAGGTCCCGACACCAAGAGAGGC 300
Qy 301 CTGGACAAGATCGAGGAGGAGCAGACAAGTGGCCAGCAGAGATCCAGCAGGGCGAGGC 360
Db 301 CTGGACAAGATCGAGGAGGAGCAGACAAGTGGCCAGCAGAGATCCAGCAGGGCGAGGC 360
Qy 361 GCCACAGGCGCAGCTGAGCCAGCACTACCCCATCTGCAGACCTGGAGGGCCAGATG 420
Db 361 GCCACAGGCGCAGCTGAGCCAGCACTACCCCATCTGCAGACCTGGAGGGCCAGATG 420
Qy 421 GTGCACGAGGCATCAGCCCGCCGACCTGAAACGCTGGGTGAAGTGTATCGAGGAGAAG 480
Db 421 GTGCACGAGGCATCAGCCCGCCGACCTGAAACGCTGGGTGAAGTGTATCGAGGAGAAG 480
Qy 481 GCCTTCAGCCCGGAGGTGATCCCATGTTTACCGCCCTGAGCGAGGCGCCACCCCGAG 540
Db 481 GCCTTCAGCCCGGAGGTGATCCCATGTTTACCGCCCTGAGCGAGGCGCCACCCCGAG 540
Qy 541 GACCTGACACGATGTTGAACACCGTGGGCGGCCACAGCGCCCATCGACATGCTGAAG 600
Db 541 GACCTGACACGATGTTGAACACCGTGGGCGGCCACAGCGCCCATCGACATGCTGAAG 600
Qy 601 GACACCATCAACGAGGAGCGCCGAGTGGGACCGCGTGACACCCCTGCAGCGCGGCC 660
Db 601 GACACCATCAACGAGGAGCGCCGAGTGGGACCGCGTGACACCCCTGCAGCGCGGCC 660
Qy 661 ATCGCCCGCGCCAGATGCGGAGCGCCCGCGGCGAGCAGATGCGCGGCGACCAACGACCC 720
Db 661 ATCGCCCGCGCCAGATGCGGAGCGCCCGCGGCGAGCAGATGCGCGGCGACCAACGACCC 720
Qy 721 CTGCAGAGCAGATGCGCTGATGACAGCAAGCAAGATCGTGGGAGTGAAGAGTAC 780
Db 721 CTGCAGAGCAGATGCGCTGATGACAGCAAGCAAGATCGTGGGAGTGAAGAGTAC 780
Qy 781 AAGCGGTGGATCATCTGGGCTGAACAAATGTTGGGATGTACAGCGCGGTGAGATC 840
Db 781 AAGCGGTGGATCATCTGGGCTGAACAAATGTTGGGATGTACAGCGCGGTGAGATC 840
Qy 841 CTGACATCAAGCAGGCGCCCAAGAGCCCTTCGCGGACTACGTGGACCGCTTCTTCAAG 900
Db 841 CTGACATCAAGCAGGCGCCCAAGAGCCCTTCGCGGACTACGTGGACCGCTTCTTCAAG 900
Qy 901 ACCCTGCGCGGAGCAGAGCAGCAGCAGAGGTGAAGACTGAGTACCGACACCTGCTG 960
Db 901 ACCCTGCGCGGAGCAGAGCAGCAGAGGTGAAGACTGAGTACCGACACCTGCTG 960
Qy 961 GTGCAGACGCAACCCGCTGCAAGACCATCTGTGCGGCTCTCGGCCCGCGGCGAGC 1020
Db 961 GTGCAGACGCAACCCGCTGCAAGACCATCTGTGCGGCTCTCGGCCCGCGGCGAGC 1020
Qy 1021 CTGAGGAGATGATGACCGCTTGCAGGGCTGGCGGCGCCAGCCACCAAGCGCGCTG 1080
Db 1021 CTGAGGAGATGATGACCGCTTGCAGGGCTGGCGGCGCCAGCCACCAAGCGCGCTG 1080
Qy 1081 CTGCGGAGGCGATGAGCGCCAGCCACACAGCGTGTATGTCAGAGAGCAACITCAAG 1140
Db 1081 CTGCGGAGGCGATGAGCGCCAGCCACACAGCGTGTATGTCAGAGAGCAACITCAAG 1140
```

```
Qy 1141 GGCCCCCGGCGATCGTCAAGTGTCTCAACTGCGCAAGAGGCGCCACATGCGCCGCAAC 1200
Db 1141 GGCCCCCGGCGATCGTCAAGTGTCTCAACTGCGCAAGAGGCGCCACATGCGCCGCAAC 1200
Qy 1201 TGCGGCGCCCGCGCAAGAGGCTGCTGGAAGTCCGCAAGAGGCGCCACAGATGAAG 1260
Db 1201 TGCGGCGCCCGCGCAAGAGGCTGCTGGAAGTCCGCAAGAGGCGCCACAGATGAAG 1260
Qy 1261 GACTGCACGAGCGCGCAGGCGCAACTTCTTGGCAAGATCTGGCCAGCGCACAGGCGCG 1320
Db 1261 GACTGCACGAGCGCGCAGGCGCAACTTCTTGGCAAGATCTGGCCAGCGCACAGGCGCG 1320
Qy 1321 CCGGGCAACTTCTGAGAGCGCGCGGAGCGCACTGCGCCCGCGCGCGAGAGCTTCGCG 1380
Db 1321 CCGGGCAACTTCTGAGAGCGCGCGGAGCGCACTGCGCCCGCGCGCGAGAGCTTCGCG 1380
Qy 1381 TTGAGGAGACCAACCCCGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Db 1381 TTGAGGAGACCAACCCCGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Qy 1441 CTGAAGAGCTGTTCGCGACGAGCGCCCTGACGCGAGTAA 1479
Db 1441 CTGAAGAGCTGTTCGCGACGAGCGCCCTGACGCGAGTAA 1479

RESULT 3
AX468547
LOCUS AX468547 1479 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 67 from Patent WO0226209.
ACCESSION AX468547
VERSION AX468547.1 GI:21901377
KEYWORDS Human immunodeficiency virus 1 (HIV-1)
SOURCE Human immunodeficiency virus 1
ORGANISM Human immunodeficiency virus 1
VIRUSES: Retroviridae; Retroviridae; Lentivirus; Primate
REFERENCE 1
AUTHORS O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
Singh,M., Ulmer,J. and Dubensky,T.W.
TITLE Microparticles for delivery of the heterologous nucleic acids
JOURNAL Patent: WO 0226209-A 67 04-APR-2002;
CHIRON CORPORATION (US)
FEATURES
Source 1. 1479
/organism="Human immunodeficiency virus 1"
/mol_type="genomic DNA"
/db_xref="taxon:11676"
BASE COUNT 325 a 527 c 457 g 160 t 10 others
ORIGIN
Query Match 99.6%; Score 1473.4; DB 6; Length 1479;
Best Local Similarity 99.3%; Pred. No. 1.2e-150;
Matches 1458; Conservative 10; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGCGCGCGCGCGCGAGTCTGCTGCGCGCGCGCGCGAGCTGGAGCGCGATCCGC 60
Db 1 ATGGGCGCGCGCGCGCGAGTCTGCTGCGCGCGCGCGCGAGCTGGAGCGCGATCCGC 60
Qy 61 CTGGCCCGCGGCGCAAGAGTGTACATGATGAAGCACTGTGTGGGCGACGCGGAG 120
Db 61 CTGGCCCGCGGCGCAAGAGTGTACATGATGAAGCACTGTGTGGGCGACGCGGAG 120
Qy 121 CTGGAGAAGTTGCGCCCTGAACCCCGGCTGCTGGAGACCAAGGAGGCTGCAAGCAATC 180
Db 121 CTGGAGAAGTTGCGCCCTGAACCCCGGCTGCTGGAGACCAAGGAGGCTGCAAGCAATC 180
Qy 181 ATCGCCAGCTGCAACCCGCTGTCAGACCGGAGCGAGAGCTGAAGAGCTGTTCAC 240
Db 181 ATCGCCAGCTGCAACCCGCTGTCAGACCGGAGCGAGAGCTGAAGAGCTGTTCAC 240
Qy 241 ACCGTGGCCACCTGTACTGCTGCACGAGAAGATCGAGGTCCCGACACCAAGAGAGGC 300
Db 241 ACCGTGGCCACCTGTACTGCTGCACGAGAAGATCGAGGTCCCGACACCAAGAGAGGC 300
```

```
Db 241 ACCGTGGCCACCCCTGTACTCGCTGACAGAGAATCGAGGTGCGGCACACCAAGAGGCC 300
QY 301 CTGGACAAGATCGAGGAGGAGCAGACAGTGTCCAGGACAGATCCAGCAGGCCAGGCC 360
Db 301 CTGGACAAGATCGAGGAGGAGCAGACAGTGTCCAGGACAGATCCAGCAGGCCAGGCC 360
QY 361 GCGCAAGAGGCAAGGTGAGCCAGCAACTACCCCATCTGTCAGAACTCTGAGGGTCAGATG 420
Db 361 GCGCAAGAGGCAAGGTGAGCCAGCAACTACCCCATCTGTCAGAACTCTGAGGGCCAGATG 420
QY 421 GTGCACAGGCCATCAGCCCCCGCCACCTCGAAGCGCTGGTGAAGGTGATCGAGAGAG 480
Db 421 GTGCACAGGCCATCAGCCCCCGCCACCTCGAAGCGCTGGTGAAGGTGATCGAGAGAG 480
QY 481 GCGTTACAGCCCGAGGTGATCCCATCTGTCAGGAGGAGGAGGAGGAGGAGGAGGAG 540
Db 481 GCGTTACAGCCCGAGGTGATCCCATCTGTCAGGAGGAGGAGGAGGAGGAGGAGGAG 540
QY 541 GACCTGAACACAGATGTTGAACACCGTGGCGGCCACCAAGCGCCCATCGATCGATGAAG 600
Db 541 GACCTGAACACAGATGTTGAACACCGTGGCGGCCACCAAGCGCCCATCGATCGATGAAG 600
QY 601 GACACCATCAACGAGGAGGCCCGCAGTGGGACCGGTGACACCCCGTGCAGCGCGGCC 660
Db 601 GACACCATCAACGAGGAGGCCCGCAGTGGGACCGGTGACACCCCGTGCAGCGCGGCC 660
QY 661 ATGCGCCCCCGGCAGATGGGAGGAGCGCGCGGAGGAGATCGCGCGGACCAAGCAAG 720
Db 661 ATGCGCCCCCGGCAGATGGGAGGAGCGCGCGGAGGAGATCGCGCGGACCAAGCAAG 720
QY 721 CTGCAGGACAGATCGCTTGGATGACAGCAACCCCGCCATCCCGTGGGGGACATCTAC 780
Db 721 CTGCAGGACAGATCGCTTGGATGACAGCAACCCCGCCATCCCGTGGGGGACATCTAC 780
QY 781 AAGCGGTGATCATCTCGGCGCTGAAACAGATCTGTGGGATGTACAGCCCGCGTGAAGATC 840
Db 781 AAGCGGTGATCATCTCGGCGCTGAAACAGATCTGTGGGATGTACAGCCCGCGTGAAGATC 840
QY 841 CTGGACATCAAGCAGGAGGCCCGCAAGGAGCGCTTCCCGACTAGCTGAGCGCTTCAAG 900
Db 841 CTGGACATCAAGCAGGAGGCCCGCAAGGAGCGCTTCCCGACTAGCTGAGCGCTTCAAG 900
QY 901 ACCTCTCGCGCGGAGCAGAGACACCCAGAGGTGAAGACTGATGACGACACCTCTGTG 960
Db 901 ACCTCTCGCGCGGAGCAGAGACACCCAGAGGTGAAGACTGATGACGACACCTCTGTG 960
QY 961 GTGCAGAACCGCCCAACCCCGACTGCAAGACCATCTCTGGCGCTCTCGGCGCGCGGCCAGC 1020
Db 961 GTGCAGAACCGCCCAACCCCGACTGCAAGACCATCTCTGGCGCTCTCGGCGCGCGGCCAGC 1020
QY 1021 CTGGAGGAGATGATGACCGCTGCGCAGGCGGTGGGCGGCCCGCAGCACAGGCGCGCTG 1080
Db 1021 CTGGAGGAGATGATGACCGCTGCGCAGGCGGTGGGCGGCCCGCAGCACAGGCGCGCTG 1080
QY 1081 CTGGCGGAGCGGATGAGCCAGGCGCAACACCGCTGATGATGACAGAGCAACTTCAAG 1140
Db 1081 CTGGCGGAGCGGATGAGCCAGGCGCAACACCGCTGATGATGACAGAGCAACTTCAAG 1140
QY 1141 GSCCCCGGCGCATGCTCAAGTCTTCAACTCGGCAAGGAGGCGGCGCATCGCCCGCAAC 1200
Db 1141 GSCCCCGGCGCATGCTCAAGTCTTCAACTCGGCAAGGAGGCGGCGCATCGCCCGCAAC 1200
QY 1201 TGGCGGCGCGCGCAGAGGCGTGTGGAGTGTGGCAAGGAGGCGGCGCATCGAGATGAAG 1260
Db 1201 TGGCGGCGCGCGCAGAGGCGTGTGGAGTGTGGCAAGGAGGCGGCGCATCGAGATGAAG 1260
QY 1261 GACTGACCGAGCGCGCAGCGCAACTCTCTGGGCAAGATCTGCGCCAGGCGCACAGGCGCC 1320
Db 1261 GACTGACCGAGCGCGCAGCGCAACTCTCTGGGCAAGATCTGCGCCAGGCGCACAGGCGCC 1320
QY 1321 CCGCGCAACTTCTGACAGCGCGCGCAGCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1321 CCGCGCAACTTCTGACAGCGCGCGCAGCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
```

```
QY 1381 TTCGAGGAGACACCCCGCGCCAGAACGAGGAGCAAGCAGCGGAGACCCCTGACCCAGC 1440
Db 1381 TTCGAGGAGACACCCCGCGCCAGAACGAGGAGCAAGCAGCGGAGACCCCTGACCCAGC 1440
QY 1441 CTGAAGAGCCTGTTCGGCAACGAGCCCTCTGAGCGAGTAA 1479
Db 1441 CTGAAGAGCCTGTTCGGCAACGAGCCCTCTGAGCGAGTAA 1479

RESULT 4
AX455904
LOCUS AX455904
DEFINITION Sequence 20 from Patent WO0204493.
ACCESSION AX455904
VERSION AX455904.1
KEYWORDS 1 G1:2714896
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,P.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL polypeptides and uses thereof
PATENT: WO 0204493-A 20 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
LOCATION/Qualifiers
source 1..1479
/mol_type="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="synthetic gag coding sequence of HIV strain AF110965"
BASE COUNT 325 a 533 c 461 g 160 t
ORIGIN
Query Match 98.9%; Score 1463; DB 6; Length 1479;
Best Local Similarity 99.3%; Pred. No. 1.6e-149;
Matches 1469; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 ATGGCGCCCGCGCGCAGCATCTCTCGCGCGCGGCAAGCTGGACGCTGGAGCGCATCCGC 60
Db 1 ATGGCGCCCGCGCGCAGCATCTCTCGCGCGCGGCAAGCTGGAGCGCATCCGC 60
QY 61 CTGCGCCCGCGCGCAGAAAGTGTACATGATGAAGCACCTGGTGTGGCCAGTCCCGCAG 120
Db 61 CTGCGCCCGCGCGCAGAAAGTGTACATGATGAAGCACCTGGTGTGGCCAGTCCCGCAG 120
QY 121 CTGAGAGATTTGCCCTGTAACCCCGGCTCTGTGAGACCAAGCGAGGCTGCAAGCAGATC 180
Db 121 CTGAGAGATTTGCCCTGTAACCCCGGCTCTGTGAGACCAAGCGAGGCTGCAAGCAGATC 180
QY 181 ATCCGCGAGCTGGACCCCGGCTGCAGACCGGAGGAGCTGAAGACCTGTTCAC 240
Db 181 ATCCGCGAGCTGGACCCCGGCTGCAGACCGGAGGAGCTGAAGACCTGTTCAC 240
QY 241 ACCGTGCCACCTGTACTCTCGTGCAGAGAGAAGATCGAGTCCGCGACACCAAGAGGCC 300
Db 241 ACCGTGCCACCTGTACTCTCGTGCAGAGAGAAGATCGAGTCCGCGACACCAAGAGGCC 300
QY 301 CTGGACAAGATCGAGGAGGAGCAACAAGTCCAGCAGAAAGATCCAGCAGGCGCGAGGCC 360
Db 301 CTGGACAAGATCGAGGAGGAGCAACAAGTCCAGCAGAAAGATCCAGCAGGCGCGAGGCC 360
QY 361 GCCACAAGGCGAAGGTGAGCCAGAACTACCCCATCTGTGAGAACCTGCGAGGCGCCATG 420
Db 361 GCCACAAGGCGAAGGTGAGCCAGAACTACCCCATCTGTGAGAACCTGCGAGGCGCCATG 420
QY 421 GTGCACCAAGGCGCATFAGCCCGCGCACCTCTGAACGCTGGTGAAGGTGATCGAGGAGAAG 480
Db 421 GTGCACCAAGGCGCATFAGCCCGCGCACCTCTGAACGCTGGTGAAGGTGATCGAGGAGAAG 480
QY 481 GCCTTCAGCCCCGAGGATGATCCCATCTTTCACCGCTGTGAGCGGCGCGCCAG 540
```







```
Db 1198 AACTCCGCGCCCGCCGCAAGAAGGCTGCTGGAAGTGGCGAAGGAGGCGCACAGATG 1257
QY 1258 AAGGACTGACCGAGCGCGCAGGCCAACTTCTTGGCGAGATCTGGCCAGCAGAGGGC 1317
Db 1256 AAGGACTGACCGAGCGCGCAGGCCAACTTCTTGGCGAGATCTGGCCAGCAGAGGGC 1317
QY 1318 CCCCCCGGCAACTTCTTGCAGAGCGG-----CCCGAGGCGCACCGCC 1359
Db 1318 CCCCCCGGCAACTTCTTGCAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
QY 1360 CCCCCCGGCGAGAGTTCGGCTTCGAGGAGACACCCCGCGCGAGAGCAGAGAGCAAG 1419
Db 1378 CCCCCCGGCGAGAGTTCGGCTTCGAGGAGACACCCCGCGCGCGCGCGCGCGCGCGCG 1437
QY 1420 G-----ACCGCAGAGCCCTGACCGCCCTGACAGAGCTGTCGGCAACAGCGC 1467
Db 1438 GACCGGAGCCCTACCGGAGCCCTTGACCGCCCTGCGCAGGCTGTCGGCAGCGCGCC 1497
QY 1468 CTGAGCGAGTAA 1479
Db 1498 CTGAGCGAGTAA 1509

RESULT 8
AX468544
LOCUS
DEFINITION
Sequence 64 from Patent WO0226209.
ACCESSION
AX468544
VERSION
AX468544.1 GI:21901174
KEYWORDS
Human immunodeficiency virus 1 (HIV-1);
SOURCE
Human immunodeficiency virus 1
ORGANISM
Viruses; Retroviridae; Lentivirus; Primate
Lentivirus group.
REFERENCE
1 O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
Singh,M., Uimer,J. and Dubensky,T.W.
TITLE
Microparticles for delivery of the heterologous nucleic acids
JOURNAL
Patent: WO 0226209-A 64 04-APR-2002;
CHIRON CORPORATION (US)
FEATURES
Location/Qualifiers
source
1..1509
/organism="Human immunodeficiency virus 1"
/mol_type="genomic DNA"
/db_xref="taxon:11676"
BASE COUNT 320 a 555 c 472 g 162 t
ORIGIN

Query Match 36.2% Score 1275.2; DB 6; Length 1509;
Best Local Similarity 92.5%; Pred. No. 3.2e-125;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

QY 1 ATGGGCGCGCGCGCGCAGCATCTTGGCGGCGGCAAGCTGACGCCCTGGGAGCGCATCGC 60
Db 1 ATGGGCGCGCGCGCGCAGCATCTTGGCGGCGGCAAGCTGAGCAAGTGGGAGAGATTCGC 60
QY 61 CTGCGCGCGCGCGCGCAGAAAGTGTCTACATCATGACGACACCTGTGTGGCGCGCGAG 120
Db 61 CTGCGCGCGCGCGCGCAGAAAGTGTCTACATCATGACGACACCTGTGTGTGGCGCGG 120
QY 121 CTGAGAGTTCGCCCTGAACCCCGGCTGCTGGAGACCGAGCGGCTGCAAGCAGATC 180
Db 121 CTGAGAGGCTTGGCCCTGAACCCCGGCTGCTGGAGACCGCGGCTGCAAGCAGATC 180
QY 181 ATCGCGCAGCTGACCCCGCGCTGCGAGACCGGAGCGAGCGCTGAGAGCCTGTTCAAC 240
Db 181 ATGAAGCAGCTGCGAGCGCGCGCTGCGAGACCGGAGCGAGCTGCGAGCCTGTACAC 240
QY 241 ACCGTGGCCACCCCTGTACTGCGTGCAGAGAAATCGAGTCCGCGACACCAAGAGGCC 300
Db 241 ACCGTGGCCACCCCTGTACTGCGTGCAGCGCCGCGCATCGAGTCCGCGACCAAGAGGCC 300
```

```
QY 301 CTGACCAAGATCGAGGAGGAGCAACAAGTGGCAGCAGAGATCCACGAGGCGCGAGGCC 360
Db 301 CTGACCAAGATCGAGGAGGAGCAACAAGTGGCAGCAGAGATCCACGAGGCGCGAGGAG 360
QY 361 GCCACAAAGGCGCAGGTTGAGCCCAAACTACCCATCGTGCAGAACCTGCAGGGCGCAGATG 420
Db 361 GCCAC---GGCAAGGTGAGCCAACTACCCATCGTGCAGAACCTGCAGGGCGCAGATG 417
QY 421 GTGCACAGGCGCATACGCCCGCCAGCCCTGAACCCCTGGGTGAAGGTGATCGAGGAGAAG 480
Db 418 GTGCACAGGCGCATACGCCCGCCAGCCCTGAACCCCTGGGTGAAGGTGATCGAGGAGAAG 477
QY 481 GCCTTCAGCCCGGAGGTGATGCCCATGTTTACCCCGCTGAGCGAGGCGGCGACCCCGCAG 540
Db 478 GCCTTCAGCCCGGAGGTGATGCCCATGTTTACCCCGCTGAGCGAGGCGGCGACCCCGCAG 537
QY 541 GACCTGACACGATGTTGAACACCGTGGCGCGCCAGCAGCGCCATGCGAGATCGACATCIGAAG 600
Db 538 GACCTGACACGATGTTGAACACCGTGGCGCGCCAGCAGCGCCATGCGAGATCGACATCIGAAG 597
QY 601 CACACCATCAACGAGGAGCGCGCGAGTGGACCGCGCTGCACCCCGTGCAGCGCGCGCCC 660
Db 598 CACACCATCAACGAGGAGCGCGCGAGTGGACCGCGCTGCACCCCGTGCAGCGCGCGCCC 657
QY 661 ATGCGCCCGCGCAGATCGCGAGCGCGCGCGAGTGGACCGCGCTGCACCCCGTGCAGCGCGCGCCC 720
Db 658 ATGCGCCCGCGCAGATCGCGAGCGCGCGCGAGTGGACCGCGCTGCACCCCGTGCAGCGCGCCC 717
QY 721 CTGAGGAGCAGATCGCGCTGATGACACGACCAACCCCGCTGCAGCGCGCGCATCTAC 780
Db 718 CTGAGGAGCAGATCGCGCTGATGACACGACCAACCCCGCTGCAGCGCGCGCATCTAC 777
QY 781 AAGCGGTGGATCATCTCGGCGCTGAACAAGATCGTGGCGATGTACAGCCCGTGAAGCATC 840
Db 778 AAGCGGTGGATCATCTCGGCGCTGAACAAGATCGTGGCGATGTACAGCCCGTGAAGCATC 837
QY 841 CTGACATCAAGCAGGCGCGCGCAAGAGCCCTTCCGCGACTAGTGGAGCGCTTCTTCAG 900
Db 838 CTGACATCAAGCAGGCGCGCGCAAGAGCCCTTCCGCGACTAGTGGAGCGCTTCTTCAG 897
QY 901 ACCCTGGCGCGCAGCAGGACCGCGAGAGTGTGAAGAACTGATGATGATGATGATGATGATG 960
Db 898 ACCCTGGCGCGCAGCAGGACCGCGAGAGTGTGAAGAACTGATGATGATGATGATGATGATG 957
QY 961 GTGCAGAACCGCAACCCCGCTGCAAGACCATCTCTCGCGGTCTCGGCGCGCGCGCGCGC 1020
Db 958 GTGCAGAACCGCAACCCCGCTGCAAGACCATCTCTCGCGGTCTCTCGGCGCGCGCGCGC 1017
QY 1021 CTGAGGAGATGATGACCGCTGCGCAGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Db 1018 CTGAGGAGATGATGACCGCTGCGCAGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCG 1077
QY 1081 CTGGCGAGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137
Db 1078 CTGGCGAGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137
QY 1138 AAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1197
Db 1138 AAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1197
QY 1198 AACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1257
Db 1198 AACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1257
QY 1258 AAGGACTGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1317
Db 1258 AAGGACTGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1317
QY 1318 CCCCCCGGCAACTTCTTGCAGAGCGG-----CCCGAGGCGCGCACCGCC 1359
Db 1318 CCCCCCGGCAACTTCTTGCAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
QY 1360 CCCCCCGGCGAGAGTTCGGCTTCGAGGAGACACCCCGCGCGAGAGCAGAGAGCAAG 1419
Db 1358 CCCCCCGGCGAGAGTTCGGCTTCGAGGAGACACCCCGCGCGCGCGCGCGCGCGCGCG 1417
QY 1420 G-----ACCGCAGAGCCCTGACCGCCCTGACAGAGCTGTCGGCAACAGCGC 1467
Db 1418 GACCGGAGCCCTACCGGAGCCCTTGACCGCCCTGCGCAGGCTGTCGGCAGCGCGCC 1497
QY 1468 CTGAGCGAGTAA 1479
Db 1498 CTGAGCGAGTAA 1509
```





```

RESULT 10
AX455905
LOCUS AX455905 1509 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 21 from Patent WO0204493.
ACCESSION AX455905
VERSION AX455905.1 GI:21714897
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
polypeptides and uses thereof
JOURNAL Patent: WO 0204493-A 21 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch. (ZA)
FEATURES
source
location/Qualifiers
1..1509
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="synthetic Gag coding sequence of HIV strain
AF110967"
BASE COUNT 321 a 559 c 471 g 158 t
ORIGIN
Query Match 85.2% Score 1260.8; DB 6; Length 1509;
Best Local Similarity 91.9%; Pred. No. 1.2e-127;
Matches 1389; Conservative 0; Mismatches 87; Indels 36; Gaps 4;
QY 1 ATGGGCGCGCGCCAGCATCTGCGCGGCGGCAAGCTGGAGCGCTGGAGCGCATCCGC 60
DB 1 ATGGGCGCGCGCCAGCATCTGCGCGGCGGCAAGCTGGAGCGCTGGAGCGCATCCGC 60
QY 61 CTGCGCGCGCGCGGCAAGAGTGTCTACATGATGAAGCACTGTGTGCGCGCGCGGAG 120
DB 61 CTGCGCGCGCGCGGCAAGAGTGTCTACATGATGAAGCACTGTGTGCGCGCGGAG 120
QY 121 CTGGAGAGTTCGCTGTGACCGCGCTGTGTGAGACGAGGAGGCTGTGAGAGATC 180
DB 121 CTGGAGGCTTCGCTGTGACCGCGCTGTGTGAGACGAGGAGGCTGTGAGAGATC 180
QY 181 ATCGGCGAGCTGCACCGCGCTGTGTGAGACGCGGAGGAGTGAAGAGCTGTTCAC 240
DB 181 ATGAGCAGCTGCACCGCGCTGTGTGAGACGCGGAGGAGTGAAGAGCTGTTCAC 240
QY 241 ACCGTGCGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
DB 241 ACCGTGCGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
QY 301 CTGACAAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 301 CTGACAAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 361 GCGACAAAGCGGAGGTGTGACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
DB 361 GCGAC---GCGAAGTGTGACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 421 GTGACCAAGGCTATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 418 GTGACCAAGGCTATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477
QY 481 GCCTTCAGCGCGCGGAGTGTGATCCCATGTTACCGCGCTGTGAGGAGGCGCGCGCG 540
DB 478 GCCTTCAGCGCGCGGAGTGTGATCCCATGTTACCGCGCTGTGAGGAGGCGCGCGCG 537
QY 541 GACCTGAACACGATGTGTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 538 GACCTGAACACGATGTGTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
QY 601 GACACCATCAACGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
DB 598 GACACCATCAACGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657

```

```

QY 661 ATGCGCGCGCGCGAGATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 658 GTGCGCGCGCGCGAGATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717
QY 721 CTGAGGAGCAGATGCGCTGTGATGACGAGCAACCGCGCGCGCGCGCGCGCGCGCG 780
DB 718 CTGAGGAGCAGATGCGCTGTGATGACGAGCAACCGCGCGCGCGCGCGCGCGCGCG 777
QY 781 AACCGGTGGATCATCTCTGGGCTGAACAAGATCGTGGGAATGTACAGCCCGGTGAGCATC 840
DB 778 AAGCGCTGGATCATCTCTGGGCTGAACAAGATCGTGGGAATGTACAGCCCGGTGAGCATC 837
QY 841 CTGACATCAAGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 838 CTGACATCCCGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897
QY 901 ACCCTGCGCGCGCGAGCAGAGACCCAGAGGTGAAGAACTGGATGACCGACACCCCTGCTG 960
DB 898 ACCCTGCGCGCGCGAGCAGAGACCCAGAGGTGAAGAACTGGATGACCGAGACCCCTGCTG 957
QY 961 GTGAGAACGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
DB 958 GTGAGAACGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1017
QY 1021 CTGAGGAGATGATGACCGCTGCCAGGGCGTGGGCGCGCGCGCGCGCGCGCGCGCGCG 1080
DB 1018 CTGAGGAGATGATGACCGCTGCCAGGGCGTGGGCGCGCGCGCGCGCGCGCGCGCGCG 1077
QY 1081 CTGCGCGCGCGGAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137
DB 1078 CTGCGCGCGCGGAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137
QY 1138 AAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1197
DB 1138 AAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1197
QY 1198 AACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1257
DB 1198 AACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1257
QY 1258 AAGACTGCACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1317
DB 1258 AAGACTGCACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1317
QY 1318 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1359
DB 1318 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
QY 1360 CCGCGCGCGCGAGAGCTTCGCGTTCGAGGAGACACCGCGCGCGCGCGCGCGCGCGCGCG 1419
DB 1378 CCGCGCGCGCGAGAGCTTCGCGTTCGAGGAGACACCGCGCGCGCGCGCGCGCGCGCG 1437
QY 1420 G-----ACCGCGAGACCGCTGACGAGCGCTGAAGAGCGCTGTTGCGGCAAGGACCG 1467
DB 1438 GACCGCGAGCGCTACCGCGAGCGCGCTGACGCGCGCTGCGCGCGCTGTTGCGGCAAGGACCG 1497
QY 1458 CTGAGCGCGTAA 1479
DB 1498 CTGAGCGCGTAA 1509

```

```

RESULT 11
AX427930
LOCUS AX427930 9166 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 168 from Patent WO0232943.
ACCESSION AX427930
VERSION AX427930.1 GI:21538017
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1

```

AUTHORS Huang, Y. and Nabel, G. J.  
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization:  
JOURNAL Patent: WO 0232943-A 168 25-APR-2002;  
GOVERNMENT OF THE UNITED STATES (US)  
FEATURES Location/Qualifiers  
Source 1. 9166  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="Plasmid pVR1012x/s containing HIV genes"  
BASE COUNT 2227 a 2662 c 2513 g 1764 t  
ORIGIN

Query Match 82.6%; Score 1221.2; DB 6; Length 9166;  
Best Local Similarity 90.1%; Pred. No. 1.3e-123;  
Matches 1346; Conservative 0; Mismatches 133; Indels 15; Gaps 3;

QY 1 ATGGGCGCGCGCCAGCATCTCTGGCGCGCAAGCTGGAGCGCTGGAGCGCATCGC 50  
DB 1880 ATGGGCGCGCGCGCGCATCTCTGGCGCGCAAGCTGGAGCGCTGGAGCGCATCAAG 1339  
QY 61 CTGCGCGCGCGCGCGCAAGTGTACATGATGAAGCACTGTGTGGGCGACCGCGAG 120  
DB 1940 CTGAAGCGCGCGCGCAAGCACTACATGATGAAGCACTGTGTGGGCGACCGCGAG 1959  
QY 121 CTGGAAGATTCGCGCTTAACCGCGCGCTGTGGAGACCGAGGCTGCAAGCATC 180  
DB 2000 CTGGAGCGCTTCGCGCGTGGACCGCGCGTGTGGAGACCGAGGCTGCAAGCATC 2059  
QY 181 ATCGCGCGCGCGCGCGCGCTGCAGACCGCGCGAGGAGCTGAAGAGCTGTTCAC 240  
DB 2060 ATGAAGCACTGTCAGCGCGCGCTGCAGACCGCGCGAGGAGCTGATCAGCTGCACAC 2119  
QY 241 ACCGTGGCGCGCGCTACTGTGTGCAGAGAGATCGAGGTCCGCGACACAGAGAGCC 300  
DB 2120 ACCGTGGCGCGCGCTACTGTGTGCAGAGAGATCGAGGTCCGCGACACAGAGAGCC 2179  
QY 301 CTGGACAAGTCTGAGGAGGAGCAACAAGTGCAGAGAGATCCAGGAGCGCGAGGCC 360  
DB 2180 CTGGACAAGATCAGGAGGAGGAGCAACAAGAGGAGGAGGAGGAGGAGGAGGAGG 2339  
QY 361 GCGGACAAGGCAAGTGTGAGGAGCAACTACCCCTATCGTGCAGAACTCTGCAGGCGAGATG 420  
DB 2240 GCGGACAAGGCAAGTGTGAGGAGCAACTACCCCTATCGTGCAGAACTCTGCAGGCGAGATG 2239  
QY 421 GTGCACCAAGGCGCATAGCGCGCGCGCGCGCTGAACGCTGGGTGAAGGTGATCGAGGAGAG 480  
DB 2300 GTGCACCAAGGCGCATAGCGCGCGCGCGCGCGCTGAACGCTGGGTGAAGGTGATCGAGGAGAG 2359  
QY 481 GCCTTCAGCGCGCGAGGTGATCCCGCATGTTCACCGCGCTGAGCGAGGCGCGCACCGCGCAG 540  
DB 2360 GCCTTCAGCGCGCGAGGTGATCCCGCATGTTCACCGCGCTGAGCGAGGCGCGCACCGCGCAG 2419  
QY 541 GACCTTGACACAGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
DB 2420 GACCTTGACACAGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2479  
QY 601 GACACCATCAAGGAGGAGCGCGAGTGGACCGCGGTGCACCGCGTGCACCGCGCGCGCGCGCG 660  
DB 2480 GACACCATCAAGGAGGAGCGCGAGTGGACCGCGGTGCACCGCGTGCACCGCGCGCGCGCGCG 2539  
QY 661 ATCGCGCGCGCGCGAGATCGGAGCG 720  
DB 2540 GTGGCGCGCGCGCGAGATCGGAGCG 2599  
QY 721 CTGCGAGGAGCAGATCGCTGTGATGACAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780  
DB 2600 CTGCGAGGAGCAGATCGCTGTGATGACCAACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2659  
QY 781 AAGCGGTGATCATCTCTGGCGCTGAACAAGATCGTGGCGATGTACAGCGCGTGGAGATC 840  
DB 2660 AAGCGGTGATCATCTCTGGCGCTGNACAGATCGTGGCGATGTACAGCGCGCGTGGAGATC 2719

QY 841 CTGGACATCAAGCAGGCG 900  
DB 2720 CTGGACATCAAGCAGGCG 2779  
QY 901 ACCCTGCGCGCGCGAGCAGACCGCGAGGAGTGAAGAACTGGATGATGATGATGATGATGATGATG 960  
DB 2780 ACCCTGCGCGCGCGAGCAGGCG 2839  
QY 961 GTGCAGACCG 1020  
DB 2840 GTGCAGACCG 2899  
QY 1021 CTGGAGGAGATGATGACCGCGCTGCCAGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080  
DB 2900 CTGGAGGAGATGATGACCGCGCTGCCAGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2959  
QY 1081 CTGGCGGAGGCGATGAGCGCAGG---CCAAACACCGCGTGTGATGATGATGATGATGATGATGATG 1137  
DB 2960 CTGGCGGAGGCGATGAGCGAAGGAGAACACACACATCATGATGATGATGATGATGATGATGATG 3019  
QY 1138 AAGGCG 1197  
DB 3020 AAGGCG 3079  
QY 1198 AACTGCG 1257  
DB 3080 AACTGCG 3139  
QY 1258 AAGGATGTCACCGAGCG 1317  
DB 3140 AAGGATGTCACCGAGCG 3199  
QY 1318 GCG 1377  
DB 3200 AGCGCGAGGAAATTTCTTCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3259  
QY 1378 CGCTTCAGGA-----GACCACCG 1425  
DB 3260 AGCTTGGGGAAGAGCAACACTCTCTCTCAAGAGCAGGAGCGCGCGCGCGCGCGCGCGCGCGCG 3319  
QY 1426 GAGACCGCTGACCGCGCTGAAGAGCGCTGTTCGCAACAGCGCGCGCGCGCGCGCGCGCGCGCG 1479  
DB 3320 TATCTTTAGCTTCCTCAGATCACTCTTTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3373

## RESULT 12

AX427936  
LOCUS AX427936 9788 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 174 from Patent WO0232943.  
ACCESSION AX427936  
VERSION AX427936.1 GI:21538023

KEYWORDS synthetic construct  
artificial sequences.

SOURCE synthetic construct  
synthetic construct  
artificial sequences.

ORGANISM Huang, Y. and Nabel, G. J.

AUTHORS Modifications of hiv env, gag, and pol enhance immunogenicity for

TITLE genetic immunization

JOURNAL Patent: WO 0232943-A 174 25-APR-2002;

GOVERNMENT OF THE UNITED STATES (US)

Location/Qualifiers

1. .9788

/organism="synthetic construct"

/mol\_type="genomic DNA"

/db\_xref="taxon:32630"

/note="Plasmid pVR1012x/s containing HIV genes"

BASE COUNT 2377 a 2817 c 2695 g 1899 t

ORIGIN

Query Match 81.5%; Score 1206; DB 6; Length 9788;

Best Local Similarity 90.0%; Pred. No. 5.8e-122;

QY	1061	CTGGCCGAGGGATGAGCCAGG----	CAAACACACACCGGTATCATCTCAAGAGCAACTTC	1117						
DB	2960	CTGGCGGAGGCATGAGCAAGGTGAACAACACACCAACATCATGATSCAGCGGAGCAACTGC	3019							
QY	1138	AAGGCCCGCGCCCATCGTCAAGTGCITCAACTGGGCAAGAGGSCCAATGCGCCGC	1157							
DB	3020	AAGGCCCGCAAGCGACCATCAAGTGCITCAACTGGCGCAAGAGGGCCACTGGCCCGC	3079							
QY	1198	AACTGCGCGCCCGCCCAAGAAGGCGTGTGGAACTGCGGCAAGAGGGCCACCAAGT	1257							
DB	3080	AACTGCGCGCCCGCCCGCAAGAGGCGTGTGGAACTGCGGCAAGAGGGCCACCAAGT	3139							
QY	1258	AAGGACTGCACGAGCGCCAGGCCAACTCTCTGGGCAAGATCTGGCCGAGCCACAAGGC	1317							
DB	3140	AAGGACTGCACGAGCGCACAGGTTAA-----TAGGGAAGATCTGGCTTCCGACAAGGGA	3194							
QY	1318	CGCCCGGCAACTTCTGTCAGAGCGCCCGGAGCGCCAGCCGCCCGCCCGCGAGAGTTC	1377							
DB	3195	AGGCAGGGAATTTCTTCAGAGCACACAGAGCCCAACAGCCCCACCAAGAGAGTTC	3254							
QY	1378	CGCTTCGAGGA-----GACCACCCCGCGCCAGAGCAGGAG-----AGCAAGAGCGGC	1425							
DB	3255	AGGTTTGGGAAGAGACAACAACCTCCCTCTCAGAAAGCAGGAGCCGATAGCAAGGAACCTG	3314							
QY	1426	GAGACCTTGACAGCGCTGAAGAGCCTGTTCGCAAGCAGCCCTCAGGCAGTAA	1479							
DB	3315	TATCCTTTAGCTTCCCTCAGATCACTCTTGGCAGCGACCCCTCGTCACATTA	3368							
RESULT 13										
AX149648										
LOCUS	AX149648	4288 bp	DNA	linear						
DEFINITION	Sequence 2 from Patent WO0136614.									
ACCESSION	AX149648									
VERSION	AX149648.1	GI:14348047								
KEYWORDS	Human immunodeficiency virus									
SOURCE	Human immunodeficiency virus									
ORGANISM	Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate lentivirus group.									
REFERENCE	1									
AUTHORS	Shao, Y., Wagner, R., Wolf, H. and Graf, M.									
TITLE	The genome of the hiv-1 inter-subtype (c/b') and use thereof									
JOURNAL	Patent: WO 0136614-A 2 25-MAY-2001;									
	Gesamt: KNB Gesellschaft fuer angewandte Biotechnologie (DE) ;									
	Shao, Yiming (CN)									
FEATURES										
source										
1. 4288										
/organism="Human immunodeficiency virus"										
/mol_type="genomic DNA"										
/db_xref="taxon:12721"										
BASE COUNT	1025 a	1358 c	1397 g	507 t						
ORIGIN										
Query Match 81.5%; Score 1205.4; dB 6; Length 4288;										
Best Local Similarity 89.5%; Pred. No. 9.5e-122;										
Matches 1335; Conservative 0; Mismatches 141; Indels 15; Gaps 3;										
QY	1	ATGGCGCCCGCGCCAGCATCTCGCGGCGCGCAAGCTGGAGCCCTGGAGCGCATCGCC	60							
DB	13	ATGGCGCCAGCGGCGCATCTCGCGGCGCGCAAGCTGGAGCGCATCGCC	72							
QY	61	CTGCGCCCGCGCGGCAAGAGTGTACATGATGAAGCACTGCTGTGGCCAGCGCGAG	120							
DB	73	CTGAGCGCCCGCGGCAAGAGCACTACATGCTGAAGCACTGTGTGGCCAGCGAGG	132							
QY	121	CTGGAGAAGTTTGGCCCTGAACCCCGGCTCTCGAGACCAAGCGAGGCTGTCAAGCAGATC	180							
DB	133	CTGGAGAAGTTTGGCCCTGAACCCCGGCTCTCGAGACCAAGCGAGGCTGTCAAGCAGATC	192							
QY	181	ATCCGCCAGCTGTCACCCCGCCCTGTCAGACCGCGGAGGAGCTGAAGAGCGTGTTCAC	240							

Db 193 ATGAGACAGCTCCAGAGCGCCCTCGACAGCCGACCCAGGAGCTGAGAGCCCTGTTCCAC 252  
QY 241 ACCGTGGCAGCCTGTATCGGTGACAGAGAATCGAGGTCCGACACCAAGAGGCC 300  
Db 253 ACCGTGGCAGCCTCTACTCGGTGACAGACCAGATCGAGCTGAGGACACACAGGAGGCC 312  
QY 301 CTGGACAGATCGAGAGGACAGAACAAAGTGCACAGAGAGATCCAGCGGCCGAGGCC 360  
Db 323 CTGGACAGATCGAGAGGACAGAACAAAGATCCAGAGAGAGCCAGAGGCAAGAG 372  
QY 361 GCGCAAGAGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAAACCTTCAGAGGCGAGATG 420  
Db 373 GCGCAAGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAAACCTTCAGAGGCGCAAG 429  
QY 421 GTGCACAGGCCATCAATGCGCGGACCTGAACGCTGGGTGAAGGTGATCGAGGAGAG 480  
Db 430 GTGCACAGGCCATCAGCCCGCAGGACCTGAATCGATGGGTGAAGGTGGTGAGAGAGAG 489  
QY 481 GCGTTCAGCCCGGAGGTGATCCCATGTTCCAGCGCTGAGCGAGGCGGCCAGGCCAG 540  
Db 490 GCGTTCAGCCCGGAGGTGATCCCATGTTCCAGCGCTGAGCGAGGCGGCCAGGCCAG 549  
QY 541 GACCTGAACAGATGTTGAACACCGTGGCGGCGACACAGCGGCCATTCACATGCTCAAG 600  
Db 550 GACCTGAACAGATGTTGAACACCGTGGCGGCGACACAGCGGCCATTCACATGCTCAAG 609  
QY 601 GACACCATCAAGAGAGGCGCGCGAGTGGACCGCTGCACCGCGTGCAGCGCGGCC 660  
Db 610 GACACCATCAAGAGAGGCGCGCGAGTGGACCGCTGCACCGCGTGCAGCGCGGCC 669  
QY 661 ATCGCCCGCGGACATCGGAGCGCGCGCGAGCGGACATCGCGCGCACCAAGAGAG 720  
Db 670 ATCGCCCGCGGACATCGGAGCGCGCGAGCGGACATCGCGCGCACCAAGAGAG 729  
QY 721 CTGCAGGACATCGCTGATGACAGACACCCCGGACATCGCGCGGAGGAGATATAC 780  
Db 730 CTGCAGGACATCGCTGATGACAGACACCCCGGACATCGCGCGGAGGAGATATAC 789  
QY 781 AAGCGGTGATCATCTCGCGGCTGAACAGATCGTGGGATGTACAGCCCGGTGAGCATC 840  
Db 790 AAGCGGTGATCATCTCGGTGAACAGATCGTGGGATGTACAGCCCGGTGAGCATC 849  
QY 841 CTGGACATCAAGAGGCGCGCAAGAGCGCTTCGCGGACTAGCTGGAGCGCTTCTCAAG 900  
Db 850 CTGGACATCAAGAGGCGCGCAAGAGCGCTTCAGGAGCTAGCTGGAGCGCTTCTCAAG 909  
QY 901 ACCCTGGCGCGGACAGAGGACGAGGAGTGAAGACTGGATGACGACGACACCGCTGCG 960  
Db 910 ACCCTGGCGCGGACAGAGGACGAGGAGTGAAGACTGGATGACGACGACACCGCTGCG 969  
QY 961 GTGCAGAGCGCAACCGGACTGCAAGACCATCTGCGCGCTCTCGGCGCGGCGGCGAGC 1020  
Db 970 GTGCAGAGCGCAACCGGACTGCAAGACCATCTGAGGCGCTTGGCGCGCGGCGGAGC 1029  
QY 1021 CTGGAGGAGATGATACCGCTGCGCAGGCGTGGGCGCGCGGCGGCGGCGGCGGCGG 1080  
Db 1030 ATCGAGGAGATGATACCGCTGCGCAGGCGTGGGCGCGCGGCGGCGGCGGCGGCGG 1089  
QY 1081 CTGGCGCGGCGGATGAGCGCGGCAACACCGCGTGTATGATGCAAGAGCAACTTCAAG 1140  
Db 1090 CTGGCGCGGCGGATGAGCGCGGCAACACCGCGTGTATGATGCAAGAGCAACTTCAAG 1149  
QY 1141 GCGCGCGCGGATGCTTAAGTGTTCATCTGCGGCAAGAGGCGGCGGCGGCGGCGGCGG 1200  
Db 1150 GCGCGCGGCGGATGCTTAAGTGTTCATCTGCGGCAAGAGGCGGCGGCGGCGGCGGCGG 1209  
QY 1201 TCGCGCGCGCGCGGCAAGAGGCGTGGAGTGGCGGAGGAGGCGGCGGCGGCGGCGGCGG 1260  
Db 1210 TCGCGCGCGCGCGGCAAGAGGCGTGGAGTGGCGGAGGAGGCGGCGGCGGCGGCGGCGG 1269  
QY 1261 GACTGACCGGCGGCGGCGGCAAGTCTCGTGGCGCAAGATCTGCGCGGCGGCGGCGGCGG 1320  
Db 1270 GACTGACCGGCGGCGGCGGCAAGTCTCGTGGCGCAAGATCTGCGCGGCGGCGGCGGCGG 1329

QY 1321 CCGCGCAACTCTCTCAGAGGCGCGCCGAGCGCCAGCCGCCCGCCCGCGGAGAGCTTCGCG 1380  
Db 1330 CCGCGCAACTCTCTCAGAGGCGCGCCGAGCGCCAGCCGCCCGCCCGCGGAGAGCTTCGAG 1389  
QY 1381 TTCGAGGAG-----ACCACCCCGCGGAGGAGGAG-----ASCAAGGACCGGAG 1428  
Db 1390 TTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1443  
QY 1429 ACCCTGACGAGCTGAGAGGCTGTTCGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1479  
Db 1450 CCCTGACGAGCTGAGAGGCTGTTCGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500

RESULT 14  
AF201927  
LOCUS Synthetic construct gag protein gene, complete cds. SYN 16-MAR-2000  
DEFINITION AF201927  
ACCESSION AF201927  
VERSION AF201927.1 GI:7248702  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
ARTIFICIAL SEQUENCES  
REFERENCE 1 (bases 1 to 1509)  
AUTHORS zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,  
Sciby,M., Otten,S.R. and Barnett,S.W.  
TITLE Increased expression and immunogenicity of sequence-modified human immunodeficiency virus type 1 gag gene  
JOURNAL 20148954  
MEDLINE 10694277  
PUBMED  
REFERENCE 2 (bases 1 to 1509)  
AUTHORS zur Megede,J. and Barnett,S.W.  
TITLE Direct Submission  
JOURNAL Submitted (04-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton,  
Emeryville, CA 94608, USA  
FEATURES  
Location/Qualifiers  
1..1509  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
1..1509  
/note="HIV-1SF2 p55gag"  
/codon\_start=1  
/transl\_table=1  
/product="gag protein"  
/protein\_id="AF43628.1"  
/db\_xref="GI:7248703"  
/translation="MGARASVLSGELDKWEKIRFGGKKKKYKHKHIVASRELERF  
AVNQLTSEGCROILQQLPSLOTGSEELSLYNTVATLYCVHQHIDKDKTKEALE  
KTEEQNKSKKKAQAAAGTGNSSQVSNYPVQNLQGMVHQAIKPRTLNAAWKV  
VKEKAFSEVTPMFSALSEGATPDNLNTMLNTVGHQAOMLKEITNEEAAMDRVH  
PVYAPIAGQMREPRGSDIAGTSTLOEIGWMTNNEPVPVETIKRYITLGLNVLV  
RWYSPSTILDTRQPKPEPRDYVDHFYKTLRAQASODVKNMETILLVONANPDKT  
ILKALGPATLEEMWTACGCGGCGGKHARVLAEMSOVNTPATIMQGRNFRKRYV  
KCFNGKEHTARNCRPFKRCNCRGHEHMDKCTPRANFLGKTPSTGRPNF  
LQSRPEPTAPPEESPRFGKEKTPPSQKQEPIDKELYPLTSLRSLFGNDPSSQ"

BASE COUNT 328 a 543 c 479 g 159 t  
ORIGIN

Query Match 81.0%; Score 1197.6; DB 12; Length 1509;  
Residual Similarity 89.3%; Pred. No. 7.6e-121;  
Matches 1347; Conservative 0; Mismatches 133; Indels 30; Gaps 4;

QY 1 ATGGGCGCGCGCGCGGAGGATCTCTGCGCGCGGCGGCAAGCTGGAGCGCTGGGAGCGCATCGCG 60  
Db 1 ATGGGCGCGCGCGCGCGGAGGATCTCTGAGCGCGCGGCGGAGCTGGAGCAAGCTGGAGAGATCGCG 60  
QY 61 CTGCGCGCGCGCGCGGCAAGAGTGTCTACATGATGAAGACCTGGTGTGGCGCGCGCGGAG 120  
Db 61 CTGCGCGCGCGCGGCAAGAGTGTCTACATGATGAAGAGTGTGAAGTGTGAAGTGTGGCGCGCGGAG 120



Qy		I	ATGGGCGCCGCCAGCATCTTCGCGGGGGAAGCTGACGCCTTGGAGAGCAATCCGC	60
Db				
Qy		37	ATGGGCGCAGGCGACGGTGCTGAGCGCGCGGAGCTGCACAAGTGGGAGAAGATCAGG	96
Db				
Qy		61	CTGGCGCCGGCGGAGAGACTCTACATGATCAAGCACTGTGTGTGGCCAGCCCGGAS	120
Db				
Qy		97	CTGAGCGCCGGCGGAGAAAGCAAGTAGTACAGCTGAAGCACATGTTGTGGCGCAGCAGGAS	156
Db				
Qy		121	CTGGAAGAATTCCGCTTGAAACCCCGGCTGCTGGAGACAGGAGGCTGCAAGCAGATC	186
Db				
Qy		157	CTGGAGAGTTCCGCTTGGAACCCCGGCTGCTGSGAGCAGCAGGAGTCCAGGCAATC	216
Db				
Qy		:81	ATCCGACAGTGCACCCCGGCTGAGACAGCGGAGAGAGAGTGAAGAGCTGTTCAGC	240
Db				
Qy		217	CTGGGCGAGTGGAGCGGAGCCTGCGAGCGGAGCGGAGGAGTGAAGAGCTGACAAAC	276
Db				
Qy		241	ACCGTGGCAACCCTGTACTGCTGCAGAGMAGATGAGTCCGCGACACCAAGAGAGGC	300
Db				
Qy		277	ACGTTGCCACCCCTGTACTGCTGTCACCAAGATGAGGTGAAGAACCAAGAGAGGC	336
Db				
Qy		301	CTGGACAAATCAGAGGAGCAGACAACTAGTCCACAGAAATCCACAGC-----GQC	354
Db				
Qy		337	CTGGAGAAGATCAGGAGGAGCAGACAGAGCAAGAAAGAAAGGCCAGCAGGSCCGCGAC	396
Db				
Qy		355	GAGCGCCGCACAAAGGCAAGGTGAGCCAGAACTACCCATCTGTGCAAGAACTTGCAGGCC	414
Db				
Qy		397	GACACGGCAACAGCCAGGTCAGTGCACCACTACCCATCTGTGCAAGAACTTGCAGGCC	456
Db				
Qy		415	CAGATGTCACACAGGCCATCAGGCCGCGCACCTTGAACGCTGGGTGAAGTGTATCGAG	474
Db				
Qy		457	CAGATGTGACACAGGCCATCAGGCCGCGCACCTTGAACGCTGGGTGAAGTGTATCGAG	516
Db				
Qy		475	GAGAAGGCCCTTACGCGCGGAGGTGATCCCACTGTTCACGCGCCTGAGCAGGCGGCCAOC	534
Db				
Qy		517	GAGAAGGCCCTTACGCGCGGAGGTGATCCCACTGTTCACGCGCCTGAGCAGGCGGCCAOC	576
Db				
Qy		535	CCCAGAGCCTGAACAGCATGTTGACACACCTGTGGCGCGCACCCAGCGCCCATTCAGATG	594
Db				
Qy		577	CCCAGAGCCTGAACACCATGCTGAACACCTGTGGCGCGCACCCAGCGCCCATTCAGATG	636
Db				
Qy		595	CTGAAGGACACCATCAACGAGGAGCGCGGAGTGGACCGCTGCACCCCTGCACGCC	654
Db				
Qy		637	CTGAAGGACACCATCAACGAGGAGCGCGGAGTGGACCGCTGCACCCCTGCACGCC	696
Db				
Qy		655	GGCCCCATGCCCGCGCCAGATTCGCGAGCCCGCGGCGAGCAGCATCGCGGCAACCC	714
Db				
Qy		697	GGCCCCATGCCCGCGCCAGATTCGCGAGCCCGCGGCGAGCAGCATCGCGGCAACCC	756
Db				
Qy		715	AGCACCTGCAGGAGCAGATCGCTTGATCACCAGCAACCCCGCATCCCGTGGCGAC	774
Db				
Qy		757	AGCACCTGCAGGAGCAGATCGCTTGATCACCAGCAACCCCGCATCCCGTGGCGAG	816
Db				
Qy		775	ATCTCAAGCGGTGGATCATCTCTGGGCTTGAACAAGATCTGTGGGATGTACAGCCCGCTG	834
Db				
Qy		817	AITCTCAAGAGTGGATCATCTCTGGGCTTGAACAAGATCTGTGGGATGTACAGCCCGCTG	876
Db				
Qy		835	AGCATCTGACATCAAGCAGGCGCCCAAGAGCCCTTCCGCGCAGTACGTGAGCCGCTTC	894
Db				
Qy		877	AGCATCTGACATCAAGCAGGCGCCCAAGAGCCCTTCCGCGCAGTACGTGAGCCGCTTC	936
Db				
Qy		895	TTCAGACCCCTTGGCGCGGAGCAGACCCAGAGAGTGAAGAACTGGATGAGCAGCAACC	954
Db				
Qy		937	TACAAAGCCCTGAGGCGCGGAGCAGCCAGCAGGTGAAGAATCGATGATACCGAGCAACC	996
Db				
Qy		955</		

Genome version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:10:43 : Search time 144.5 Seconds  
(without alignments)  
1021.419 Million cell updates/sec

Title: us-09-475-704A-2

Perfect score: 60

Sequence: 1 qacatcgcaagacccctaa.....tggacattcttcagagacc 60

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1560708 seqs. 1229959015 residues

Total number of hits satisfying chosen parameters: 1177748

Minimum DB seq length: 6

Maximum DB seq length: 160

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US05\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed. and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	DB Length	ID	Description
C 1	21	35.0	87	9	US-09-864-761-28254
2	20.4	34.0	60	9	US-09-007-093-7
3	20.4	34.0	65	12	US-09-908-975-25099
4	20.2	33.7	65	12	US-09-908-975-35099
C 5	19.8	33.0	77	11	US-09-805-353-16
C 6	19.4	32.3	100	10	US-09-728-445-336
7	18.6	31.0	60	12	US-09-908-975-3744
8	18.4	30.7	60	12	US-09-908-975-13339
9	18.2	30.3	60	12	US-09-908-975-7983
C 10	18.2	30.3	88	12	US-10-133-973-102
C 11	18.2	30.3	89	10	US-09-971-980-18
C 12	18	30.0	65	12	US-09-908-975-32263
13	17.8	29.7	56	9	US-09-967-013-81
C 14	17.8	29.7	56	9	US-09-967-013-82
15	17.8	29.7	56	11	US-09-863-733A-80
C 16	17.8	29.7	56	11	US-09-863-733A-81

17	17.8	29.7	56	14	US-10-116-420-80	Sequence 80, Appl
C 18	17.8	29.7	56	14	US-10-116-420-81	Sequence 81, Appl
19	17.8	29.7	59	9	US-09-967-013-88	Sequence 88, Appl
C 20	17.8	29.7	59	9	US-09-967-013-89	Sequence 89, Appl
21	17.8	29.7	59	11	US-09-863-733A-87	Sequence 87, Appl
C 22	17.8	29.7	59	11	US-09-863-733A-88	Sequence 88, Appl
23	17.8	29.7	59	11	US-09-863-733A-93	Sequence 93, Appl
C 24	17.8	29.7	59	11	US-09-863-733A-94	Sequence 94, Appl
25	17.8	29.7	59	14	US-10-116-420-87	Sequence 87, Appl
C 26	17.8	29.7	59	14	US-10-116-420-93	Sequence 93, Appl
27	17.8	29.7	59	14	US-10-116-420-93	Sequence 93, Appl
C 28	17.8	29.7	59	14	US-10-116-420-94	Sequence 94, Appl
C 29	17.8	29.7	60	12	US-09-908-975-13021	Sequence 13021, A
30	17.8	29.7	60	12	US-09-908-975-13510	Sequence 13510, A
31	17.8	29.7	60	12	US-09-908-975-16956	Sequence 16956, A
32	17.8	29.7	60	12	US-09-908-975-16956	Sequence 16956, A
33	17.8	29.7	63	9	US-09-811-672-4	Sequence 20072, A
34	17.8	29.7	63	12	US-10-161-660-4	Sequence 4, Appl
C 35	17.8	29.7	65	12	US-09-908-975-1545	Sequence 1045, Ap
36	17.8	29.7	65	12	US-09-908-975-24242	Sequence 24242, A
37	17.8	29.7	90	10	US-09-971-788-6	Sequence 6, Appl
C 38	17.6	29.3	50	12	US-10-322-360-119	Sequence 119, App
C 39	17.6	29.3	56	9	US-09-760-574-10	Sequence 10, App
C 40	17.6	29.3	56	11	US-09-766-442A-10	Sequence 10, Appl
41	17.6	29.3	63	12	US-09-908-975-8727	Sequence 8727, Ap
C 42	17.6	29.3	85	9	US-09-815-242-1201	Sequence 1201, Ap
C 43	17.4	29.0	30	10	US-09-777-430A-21	Sequence 21, Appl
C 44	17.4	29.0	30	11	US-09-864-636A-296	Sequence 296, App
C 45	17.4	29.0	30	12	US-09-758-282-164	Sequence 164, App

#### ALIGNMENTS

#### RESULT 1

US-09-864-761-28254/c  
: Sequence 28254, Application US/U9864761  
: Patent No. US20020048763A1  
: GENERAL INFORMATION:  
: APPLICANT: Penn. Sharron G.  
: APPLICANT: Rank, David R.  
: APPLICANT: Chen, Wensheng  
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL  
: FILE REFERENCE: Aomic-X-1  
: CURRENT FILING DATE: 2001-05-23  
: PRIOR FILING DATE: 2001-05-23  
: PRIOR APPLICATION NUMBER: US 60/180,312  
: PRIOR FILING DATE: 2000-02-04  
: PRIOR APPLICATION NUMBER: US 60/207,456  
: PRIOR FILING DATE: 2000-05-26  
: PRIOR APPLICATION NUMBER: US 09/632,366  
: PRIOR FILING DATE: 2000-08-03  
: PRIOR APPLICATION NUMBER: GB 24263.6  
: PRIOR FILING DATE: 2000-10-04  
: PRIOR APPLICATION NUMBER: US 60/236,359  
: PRIOR FILING DATE: 2000-09-27  
: PRIOR APPLICATION NUMBER: PCT/US01/00666  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00667  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00664  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00669  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00665  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00668  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00663  
: PRIOR FILING DATE: 2001-01-30  
: PRIOR APPLICATION NUMBER: PCT/US01/00662



```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 69/638,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 99/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28254
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004870.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: P75080, EVALUATION 5.60e-00
US-09-864-762-28254

Query Match 35.0%; Score 21; DB 9; Length 87;
Best Local Similarity 73.0%; Pred. No. 3.9e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 20 AGAGCGCCCTCCGCGACTACGTCGACCGCTTCTTCAAA 56
||||||| || ||||| ||||| ||||| |||||
Db 86 AGAGCCCTCCACACATACCTGACACCATGATCAA 50

RESULT 2
US-09-007-093-7
; Sequence 7, Application US/09007093
; Patent No. US20020025315A1
; GENERAL INFORMATION:
; APPLICANT: Anand, Naveen N
; APPLICANT: Barber, Brian H
; APPLICANT: Bates, George A
; APPLICANT: Caterini, Judith E
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DELIVERY OF
; TITLE OF INVENTION: ANTIGENS TO SELECTED CELLS OF THE IMMUNE SYSTEM
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/007,093
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,576
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-765

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-007-093-7

Query Match 34.0%; Score 20.4; DB 9; Length 60;
Best Local Similarity 65.2%; Pred. No. 6.3e+02;
Matches 30; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 13 GCGCCCAAGAGAGCGCTTCGCGACTACGTCGACCGCTTCTTCAAGA 58
||||||| ||||| ||||| ||||| ||||| |||||
Db 13 GGTCTTAAAGAACCTTTAGAGACATATGTTGATGTTTATAAGA 58

RESULT 3
US-09-908-975-25099
; Sequence 25099, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: WASSERMAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPL
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25099
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-25099

Query Match 34.0%; Score 20.4; DB 12; Length 65;
Best Local Similarity 61.1%; Pred. No. 6.3e+02;
Matches 33; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 6 CCGCCAGGCGCCCAAGAGCGCTTCGCGACTACGTCGACCGCTTCTTCAAGAC 59
||||||| ||||| ||||| ||||| ||||| |||||
Db 2 CCGCCAGGAGACCCATGAGAGATGCTTACACCGCGCGACGATCATGTC 55

RESULT 4
US-09-908-975-30489
; Sequence 30489, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: WASSERMAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPL
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
```

; ORGANISM: Mus musculus  
US-09-728-445-336

```
Query Match      30.7%: Score 18.4; DB 12; Length 60;
Best Local Similarity 59.6%: Pred. No. 3.1e+03;
Matches 31; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 CATCGCCAGGCGCCCAAGGAGCCCTTCGGGACTAGCTGGAGGCTTC 54
DB 6 CATGACCAACTCTCAGGAGGAGCTGCACCAAGAGATCGATGACCTTC 59

RESULT 9
US-09-908-975-7983
: Sequence 7983, Application US/09908975
: Publication No. US20030165843A1
: GENERAL INFORMATION:
: APPLICANT: SHOSHAN, Avi
: APPLICANT: WASSERMAN, Alon
: APPLICANT: MINTZ, Eli
: APPLICANT: MINTZ, Liat
: TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING
: FILE REFERENCE: 36688-0005
: CURRENT APPLICATION NUMBER: US/09/908,975
: PRIOR FILING DATE: 2001-07-20
: PRIOR APPLICATION NUMBER: US 60/287,724
: PRIOR FILING DATE: 2001-05-02
: PRIOR APPLICATION NUMBER: US 60/221,607
: PRIOR FILING DATE: 2000-07-28
: NUMBER OF SEQ ID NOS: 3237
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO: 7983
: LENGTH: 60
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-908-975-7983

Query Match      30.3%: Score 18.2; DB 12; Length 60;
Best Local Similarity 58.2%: Pred. No. 3.6e+03;
Matches 32; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 5 TCGCGAGGCGCCCAAGGAGCCCTTCGGGACTAGCTGGAGGCTTC 59
DB 6 TCAGCTGGCTCAGCAAGGA:GCCGCCCAAGGAGGCGGAGCTTC 60

RESULT 10
US-10-133-973-102/c
: Sequence 102, Application US/10133973
: Publication No. US20030149254A1
: GENERAL INFORMATION:
: APPLICANT: Anderson, David
: TITLE OF INVENTION: METHODS AND COMPOSITIONS COMPRISING RENILLA GFP
: FILE REFERENCE: A-68531-3/RMS/CYO
: CURRENT APPLICATION NUMBER: US/10/133,973
: CURRENT FILING DATE: 2002-04-24
: PRIOR APPLICATION NUMBER: US 62/290,287
: PRIOR FILING DATE: 2001-05-10
: PRIOR APPLICATION NUMBER: US 09/710,058
: PRIOR FILING DATE: 2000-11-10
: NUMBER OF SEQ ID NOS: 107
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO: 102
: LENGTH: 88
: TYPE: DNA
: ORGANISM: Artificial sequence
: FEATURE:
: OTHER INFORMATION: primer
US-10-133-973-102

Query Match      30.3%: Score 18.2; DB 12; Length 88;
Best Local Similarity 58.2%: Pred. No. 3.6e+03;
Matches 32; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 3 CATCGCCAGGCGCCCAAGGAGCCCTTCGGGACTAGCTGGAGGCTTC 88
DB 17 CATTTCCACAGCCCTCCAGGAGAGCTTCAAGGCTTACGAAGA 59

RESULT 11
US-09-971-980-18/c
: Sequence 18, Application US/09971980
: Patent No. US20020164349A1
: GENERAL INFORMATION:
: APPLICANT: Weiner, David B.
: APPLICANT: Yang, Joo-Sung
: TITLE OF INVENTION: Compositions and Methods of Using Capsid Protein From Flaviviruses
: FILE REFERENCE: Upd-4105
: CURRENT APPLICATION NUMBER: US/09/971,980
: CURRENT FILING DATE: 2001-10-04
: PRIOR APPLICATION NUMBER: 60/237,885
: PRIOR FILING DATE: 2000-10-04
: NUMBER OF SEQ ID NOS: 73
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO: 18
: LENGTH: 89
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: No. US20020164349A1 Sequence
US-09-971-980-18

Query Match      30.3%: Score 18.2; DB 10; Length 89;
Best Local Similarity 61.7%: Pred. No. 3.6e+03;
Matches 29; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 3 CATCGCCAGGCGCCCAAGGAGCCCTTCGGGACTAGCTGGAGGCTTC 49
DB 83 CTTCCGCTTACCGCCATTGCCGCCACCCGCGCTGCTGGACGCT 37

RESULT 12
US-09-908-975-30263
: Sequence 30263, Application US/09908975
: Publication No. US20030165843A1
: GENERAL INFORMATION:
: APPLICANT: SHOSHAN, Avi
: APPLICANT: WASSERMAN, Alon
: APPLICANT: MINTZ, Eli
: APPLICANT: MINTZ, Liat
: APPLICANT: FAIGLER, Simchon
: TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPL
: FILE REFERENCE: 36688-0005
: CURRENT APPLICATION NUMBER: US/09/908,975
: CURRENT FILING DATE: 2001-07-20
: PRIOR APPLICATION NUMBER: US 60/287,724
: PRIOR FILING DATE: 2001-05-02
: PRIOR APPLICATION NUMBER: US 60/221,607
: PRIOR FILING DATE: 2000-07-28
: NUMBER OF SEQ ID NOS: 3237
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO: 30263
: LENGTH: 65
: TYPE: DNA
: ORGANISM: Mus musculus
US-09-908-975-30263

Query Match      30.0%: Score 18; DB 12; Length 65;
Best Local Similarity 64.3%: Pred. No. 4.2e+03;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 3 CATCGCCAGGCGCCCAAGGAGCCCTTCGGGACTAGCTGGAGGCTTC 44
DB 17 CATTTCCACAGCCCTCCAGGAGAGCTTCAAGGCTTACGAAGA 59
```

RESULT 15  
US-09-863-733A-80  
; Sequence 80, Application: US/09863733A  
; Publication No. US20030082537A1  
; GENERAL INFORMATION:  
; APPLICANT: Scantec, Jr., Vincent P.  
; TITLE OF INVENTION: METHODS FOR GENETIC ANALYSIS OF DNA  
; FILE REFERENCE: 11926-113001

Search completed: September 22, 2003, 17:30:58  
Job time : 145.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic: - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 : Search time 44 seconds  
(without alignments)  
561.886 Million cell updates/sec

Title: US-09-475-704A-2

Perfect score: 60

Sequence: 1 gacatcgccagaguccaaa.....tggacccattcttcaagac: 60

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 223691565 residues

Total number of hits satisfying chosen parameters: 330496

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents.NA: \*
- 1: /cqn2\_6/ptodata/2/ina/5A.COMB.seq: \*
  - 2: /cqn2\_6/ptodata/2/ina/5B.COMB.seq: \*
  - 3: /cqn2\_6/ptodata/2/ina/6A.COMB.seq: \*
  - 4: /cqn2\_6/ptodata/2/ina/6B.COMB.seq: \*
  - 5: /cqn2\_6/ptodata/2/ina/PCTUS.COMB.seq: \*
  - 6: /cqn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result, being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21.8	36.3	84	1	US-08-405-702A-7
2	20.8	34.7	84	1	US-08-405-702A-5
3	20.8	34.7	84	1	US-08-405-702A-9
4	20.4	34.0	60	3	US-08-943-136-7
5	20.4	34.0	60	4	US-08-973-518-7
6	20	33.3	72	1	US-08-303-275-192
7	20	33.3	72	1	US-08-303-275-193
8	19.6	32.7	57	6	5514566-17
9	18	30.0	38	3	US-09-282-147-28
10	18	30.0	72	2	US-08-707-237A-47
11	18	30.0	100	1	US-07-797-556-7
12	17.8	29.7	40	1	US-08-495-743-25
13	17.8	29.7	40	1	US-08-548-078-3
14	17.8	29.7	40	1	US-08-495-743-25
15	17.8	29.7	40	1	US-08-495-741-25
16	17.8	29.7	40	3	US-08-062-023-25
17	17.8	29.7	44	2	US-08-852D-35
18	17.8	29.7	44	3	US-08-510-646B-37
19	17.8	29.7	44	3	US-09-231-818-35
20	17.8	29.7	45	1	US-08-167-939A-14
21	17.8	29.7	45	1	US-08-567-538-14
22	17.8	29.7	63	3	US-08-750-419A-4
23	17.8	29.7	63	4	US-09-811-672-4
24	17.8	29.7	72	1	US-08-303-275-195
25	17.6	29.3	50	4	US-09-554-929-119
26	17.6	29.3	72	1	US-07-778-231B-72
27	17.6	29.3	72	1	US-07-963-321-72

28	17.6	29.3	72	1	US-08-290-641-72	Sequence 72, Appl
29	17.6	29.3	72	1	US-08-548-540-72	Sequence 72, Appl
30	17.6	29.3	72	5	PCT-US96-05809-72	Sequence 72, Appl
31	17.4	29.0	53	2	US-07-695-201B-13	Sequence 13, Appl
32	17.4	29.0	53	3	US-08-470-532-13	Sequence 13, Appl
33	17.2	28.7	48	1	US-08-151-574-14	Sequence 14, Appl
34	17.2	28.7	48	2	US-08-419-448-14	Sequence 14, Appl
35	17.2	28.7	48	4	US-09-233-510-14	Sequence 14, Appl
36	17.2	28.7	50	1	US-08-575-052-7	Sequence 7, Appl
37	17.2	28.7	50	1	US-08-614-516A-7	Sequence 7, Appl
38	17.2	28.7	50	1	US-08-770-557-7	Sequence 7, Appl
39	17.2	28.7	50	3	US-08-405-234-7	Sequence 16, Appl
40	17.2	28.7	65	3	US-09-191-852-16	Sequence 16, Appl
41	17.2	28.7	65	4	US-08-817-906-16	Sequence 16, Appl
42	17.2	28.7	65	5	PCT-US95-13376-16	Sequence 16, Appl
43	17.2	28.7	80	1	US-08-152-621-7	Sequence 7, Appl
44	17.2	28.7	80	1	US-08-406-691B-38	Sequence 38, Appl
45	17.2	28.7	80	5	PCT-US92-05035-7	Sequence 7, Appl

ALIGNMENTS

Result 1  
US-08-405-702A-7  
: Sequence 7, Application US/38405702A  
: Patent No. 5789389  
: GENERAL INFORMATION:  
: APPLICANT: Tarasewicz, Dariusz G  
: APPLICANT: Schott, Brigitte  
: APPLICANT: Holzmayer, Tatiana A.  
: APPLICANT: Roninsson, Igor B  
: TITLE OF INVENTION: BCL2 Derived Genetic Elements Associated  
: TITLE OF INVENTION: with Sensitivity to Chemotherapeutic Drugs  
: NUMBER OF SEQUENCES: 16  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Banner & Witcoff, Ltd.  
: STREET: 10 South Wacker Drive, Suite 3000  
: CITY: Chicago  
: STATE: Illinois  
: COUNTRY: USA  
: ZIP: 60606

COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/405,702A  
: FILING DATE: 17-MAR-1995  
: CLASSIFICATION: 514  
: ATTORNEY/AGENT INFORMATION:  
: NAME: No. 5789389nan, Kevin E  
: REGISTRATION NUMBER: 35,303  
: REFERENCE/DOCKET NUMBER: 95,332  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 312-715-1000  
: TELEFAX: 312-715-1234  
: TELEX: 910-221-5317  
: INFORMATION FOR SEQ ID NO: 7:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 84 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: cDNA  
US-08-405-702A-7

Query Match 36.3%; Score 21.6; DB 1; Length 84;  
Best Local Similarity 61.4%; Pred. No. 1.3e+02;  
Matches 35; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

3 CATCCGACGGCCCAAGAGCCCTCCCGGACTACGTGACCGCTTCTTCAAGAC 59

Db 15 C C T C C G C C A G G C C G C G A G A C T T C T C C G C C G C T A C C G C G C G A C T T C G C G A G A 70

RESULT 2  
US-08-405-702A-5  
Sequence 5, Application US/08405702A  
Patent No. 5789389  
GENERAL INFORMATION:  
APPLICANT: Tarasewicz, Dariusz G  
APPLICANT: Schott, Brigitte  
APPLICANT: Holzmayer, Tatiana A.  
APPLICANT: Roninson, Igor B  
TITLE OF INVENTION: BCL2 Derived Genetic Elements Associated  
TITLE OF INVENTION: With Sensitivity to Chemotherapeutic Drugs  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/405,702A  
FILING DATE: 17-MAR-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5789389can, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 95,332  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 910-221-5317  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-405-702A-5

Query Match 34.7%, Score 20.8, DB 1, Length 84:  
Best Local Similarity 60.7%, Pred. No. 2.7e+02;  
Matches 34; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
QY 3 CATCCGCCAGGCCCCCAGGAGCCCTTCCTCCGCGACTAGTGAGCGCTTCTTCAGGA 58  
DB 15 C C T C C G C C A G G C C G C G A G A C T T C C G C C G C T A C C G C G C G A C T T C G C G A G A 70

RESULT 3  
US-08-405-702A-9  
Sequence 9, Application US/08405702A  
Patent No. 5789389  
GENERAL INFORMATION:  
APPLICANT: Tarasewicz, Dariusz G  
APPLICANT: Schott, Brigitte  
APPLICANT: Holzmayer, Tatiana A.  
APPLICANT: Roninson, Igor B  
TITLE OF INVENTION: BCL2 Derived Genetic Elements Associated  
TITLE OF INVENTION: With Sensitivity to Chemotherapeutic Drugs  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/405,702A  
FILING DATE: 17-MAR-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5789389can, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 95,332  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 910-221-5317  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-405-702A-9

Query Match 34.7%, Score 20.8, DB 1, Length 84:  
Best Local Similarity 60.7%, Pred. No. 2.7e+02;  
Matches 34; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
QY 3 CATCCGCCAGGCCCCCAGGAGCCCTTCCTCCGCGACTAGTGAGCGCTTCTTCAGGA 58  
DB 15 C C T C C G C C A G G C C G C G A G A C T T C C G C C G C T A C C G C G C G A C T T C G C G A G A 70

RESULT 4  
US-08-943-136-7  
Sequence 7, Application US/08943136  
Patent No. 6291208  
GENERAL INFORMATION:  
APPLICANT: Anand, Naveen N  
APPLICANT: Barber, Brian H  
APPLICANT: Cates, George A  
APPLICANT: Caterini, Judith E  
APPLICANT: Klein, Michel H  
TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DELIVERY OF  
TITLE OF INVENTION: ANTIGENS TO SELECTED CELLS OF THE IMMUNE SYSTEM  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/943,136  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/483,576  
FILING DATE: 07-JUNE-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:

Query Match 34.7%, Score 20.8, DB 1, Length 84:  
Best Local Similarity 60.7%, Pred. No. 2.7e+02;  
Matches 34; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
QY 3 CATCCGCCAGGCCCCCAGGAGCCCTTCCTCCGCGACTAGTGAGCGCTTCTTCAGGA 58  
DB 15 C C T C C G C C A G G C C G C G A G A C T T C C G C C G C T A C C G C G C G A C T T C G C G A G A 70

```

Db      13  GCTCCATAAGAACCTTTTAGAGACIATGTGATAGATTATTAAGA 58

RESULT 6
US-08-303-275-192
: Sequence 192, Application US/08303275
: Patent No. 5766598
: GENERAL INFORMATION:
: APPLICANT: Paoletti, Enzo
: APPLICANT: Tartaglia, James
: APPLICANT: Cox, William I.
: TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS RECOMBINANT
: TITLE OF INVENTION: POXVIRUS VACCINE
: NUMBER OF SEQUENCES: 205
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Curtis, Morris & Safford
: ADDRESSEE: c/o William S. Frommer
: STREET: 530 Fifth Avenue
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/303,275
: FILING DATE:
: PRIOR CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/897,382
: FILING DATE: 11-JUN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Frommer, William S.
: REGISTRATION NUMBER: 25,506
: REFERENCE/DOCKET NUMBER: 454110-2420
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 840-3333
: TELEFAX: (212) 840-0712
: INFORMATION FOR SEQ ID NO.: 192:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 72 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-303-275-192

Query Match          33.34; Score 20; DB 1; Length 72;
Best Local Similarity 61.5%; Pred. No. 4.8e+02;
Matches 32; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY      5  TCGCCAGGGCCCCNAGGACGCTTCGGGACTACGTGGACGCCTTCTTCAA 56
       ||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db      5  TCAGAAAGGACCTTAAGACCTTTAGAGATTATAGATGAATTTATRA 56

RESULT 7
US-08-303-275-193/c
: Sequence 193, Application US/08303275
: Patent No. 5766598
: GENERAL INFORMATION:
: APPLICANT: Paoletti, Enzo
: APPLICANT: Tartaglia, James
: APPLICANT: Cox, William I.
: TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS RECOMBINANT
: TITLE OF INVENTION: POXVIRUS VACCINE
: NUMBER OF SEQUENCES: 205
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Curtis, Morris & Safford
: ADDRESSEE: c/o William S. Frommer
: STREET: 530 Fifth Avenue

```





;; FILING DATE: 29-OCT-1987  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 06/927,258  
;; FILING DATE: 04-NOV-1986  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Trecartin, Richard F.  
;; REGISTRATION NUMBER: 31,501  
;; REFERENCE/DOCKET NUMBER: A-55186-10/WHD  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 781-1989  
;; TELEFAX: (415) 398-3249  
;; TELEX: 910 277299  
;; INFORMATION FOR SEQ ID NO: 47:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 72 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: unknown  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: DNA (genomic)  
US-08-707-237A-47

Query Match 30.0%; Score 18; DB 2; Length 72;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACATCCGCCAGGCCCA 19  
Db 26 ACA:CCGCCAGGCCCA 43

RESULT 1:  
US-07-797-556-7/c  
; Sequence 7, Application US/07797556  
; Patent No. 5262522  
; GENERAL INFORMATION:  
; APPLICANT: Gearing, David P.  
; TITLE OF INVENTION: Receptor for Oncostatin M and Leukemia  
; TITLE OF INVENTION: Inhibitory Factor  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/797,556  
; FILING DATE: 19911122  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Seese, Kathryn A.  
; REGISTRATION NUMBER: 32,172  
; REFERENCE/DOCKET NUMBER: 2607  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-587-0430  
; TELEFAX: 206-587-0606  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 100 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-07-797-556-7

Query Match 30.0%; Score 18; DB 1; Length 100;  
Best Local Similarity 56.9%; Pred. No. 2.1e+03;

Matches 33; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
Qy 3 CATCCGCCAGGCCCTTCGGCACACCTGGGACCGCTTCTTCAAGACC 60  
Db 73 CCTCCACCTGATCCACCACTCTGTATCTCCACCACTGAACCACTCCCAAGAAC 16

RESULT 12  
US-08-495-743-25/c  
; Sequence 25, Application US/08495743  
; Patent No. 5634416  
; GENERAL INFORMATION:  
; APPLICANT: Thomas J. Cummins, Susan M. Atwood  
; APPLICANT: Lynn Bergmeyer, John B. Findlay  
; APPLICANT: John W.H. Sutherland, JoAnne H. Kerschner  
; TITLE OF INVENTION: DIAGNOSTIC COMPOSITIONS, ELEMENTS,  
; TITLE OF INVENTION: METHODS AND TEST KITS FOR  
; TITLE OF INVENTION: AMPLIFICATION AND DETECTION OF TWO  
; TITLE OF INVENTION: OR MORE TARGET DNA'S USING PRIMERS  
; TITLE OF INVENTION: HAVING MATCHED MELTING TEMPERATURES  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eastman Kodak Company, Patent Legal Staff  
; STREET: 343 State Street  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14650 - 2201  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5inch, 1.44 Mb storage (IBM)  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS Version 3.3  
; SOFTWARE: PC-8 (Word for Windows)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/495,743  
; FILING DATE: 08-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/062,023  
; FILING DATE: 14-MAY-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Tucker, J. Lanny  
; REGISTRATION NUMBER: 27,678  
; REFERENCE/DOCKET NUMBER: 67271A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 722-9332  
; TELEFAX: (716) 477-4646  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 nucleotides  
; TYPE: Nucleic acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
; MOLECULE TYPE: Primer for HIV-1 DNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE: Synthetically prepared  
; IMMEDIATE SOURCE: Same  
; PUBLICATION INFORMATION: Unknown  
US-08-495-743-25

Query Match 29.7%; Score 17.8; DB 1; Length 40;  
Best Local Similarity 75.9%; Pred. No. 2.3e+03;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Qy 1 GACATCGCCAGGCCGCCAGGAGCCCTT 29  
Db 32 GACATAGACAGGAGGACCAAGAACCCCTT 4

RESULT 13  
US-08-548-078-3/c  
; Sequence 3, Application US/08548078

Patent No. 5674717  
GENERAL INFORMATION:  
APPLICANT: John W. Backus  
APPLICANT: William H. Findlay  
APPLICANT: John W. H. Sutherland  
APPLICANT: Marlene M. King  
TITLE OF INVENTION: RAPID METHOD FOR DIFFERENTIAL  
TITLE OF INVENTION: COMPARISON OF TWO DIFFERENT NUCLEIC ACID  
TITLE OF INVENTION: SEQUENCES USING POLYMERASE CHAIN REACTION  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eastman Kodak Company,  
ADDRESSEE: Patent Legal Staff  
STREET: 343 State Street  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14650 - 2201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage (IBM)  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS Version 3.3  
SOFTWARE: PC-8 (Word for Windows 2.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/548,078  
FILING DATE:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Tucker, J. Lanny  
REGISTRATION NUMBER: 27,678  
REFERENCE/DOCKET NUMBER: 58258  
TELEPHONE: (716) 722-9332  
TELEFAX: (716) 477-4646  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 nucleotides  
TYPE: Nucleic acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: HIV-1 DNA primer  
HYPOTHETICAL: No  
ANTI-SENSE: No  
ORIGINAL SOURCE: Synthetically prepared  
IMMEDIATE SOURCE: Same  
PUBLICATION INFORMATION: No. 5674717e  
US-08-548-078-3

Query Match 29.7%; Score 17.8; DB 1; Length 40;  
Best Local Similarity 75.9%; Pred. No. 2.3e+03;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GACATCGCCAGGCCCCAAGGAGCCCTT 29  
Db 32 GACATAAGACAGGACCAAGGAGCCCTT 4

## RESULT 14

US-08-495-739-25/c  
Sequence 25, Application US/08495739  
Patent No. 5702901  
GENERAL INFORMATION:

APPLICANT: Thomas J. Cummins, Susan M. Atwood  
APPLICANT: Lynn Bergmeyer, John B. Findlay  
APPLICANT: John W.H. Sutherland, Joanne H. Kerschner  
TITLE OF INVENTION: DIAGNOSTIC COMPOSITIONS, ELEMENTS,  
METHODS AND TEST KITS FOR

TITLE OF INVENTION: AMPLIFICATION AND DETECTION OF TWO  
TITLE OF INVENTION: OR MORE TARGET DNA'S USING PRIMERS  
TITLE OF INVENTION: HAVING MATCHED MELTING TEMPERATURES  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eastman Kodak Company, Patent Legal Staff  
STREET: 343 State Street  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14650 - 2201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage (IBM)  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS Version 3.3  
SOFTWARE: PC-8 (Word for Windows)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/495,739  
FILING DATE: 08-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/062,023  
FILING DATE: 14 MAY 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Tucker, J. Lanny  
REGISTRATION NUMBER: 27,678  
REFERENCE/DOCKET NUMBER: 67271A  
TELEPHONE: (716) 722-9332  
TELEFAX: (716) 477-4646  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 nucleotides  
TYPE: Nucleic acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Primer for HIV-1 DNA  
HYPOTHETICAL: No  
ANTI-SENSE: No  
ORIGINAL SOURCE: Synthetically prepared  
IMMEDIATE SOURCE: Same  
PUBLICATION INFORMATION: Unknown  
US-08-495-739-25

Query Match 29.7%; Score 17.8; DB 1; Length 40;  
Best Local Similarity 75.9%; Pred. No. 2.3e+03;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GACATCGCCAGGCCCCAAGGAGCCCTT 29  
Db 32 GACATAAGACAGGACCAAGGAGCCCTT 4

## RESULT 15

US-08-495-741-25/c  
Sequence 25, Application US/08495741;  
Patent No. 5733751  
GENERAL INFORMATION:

APPLICANT: Thomas J. Cummins, Susan M. Atwood  
APPLICANT: Lynn Bergmeyer, John B. Findlay  
APPLICANT: John W.H. Sutherland, Joanne H. Kerschner  
TITLE OF INVENTION: DIAGNOSTIC COMPOSITIONS, ELEMENTS,  
METHODS AND TEST KITS FOR  
TITLE OF INVENTION: AMPLIFICATION AND DETECTION OF TWO  
TITLE OF INVENTION: OR MORE TARGET DNA'S USING PRIMERS  
TITLE OF INVENTION: HAVING MATCHED MELTING TEMPERATURES  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eastman Kodak Company, Patent Legal Staff  
STREET: 343 State Street  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.

```

; ZIP: 14650 - 2201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5inch, 1.44 MB storage (IBM)
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS Version 3.3
; SOFTWARE: PC-8 (Word for Windows)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/495,741
; FILING DATE: 08-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/062,023
; FILING DATE: 19930514
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tucker, J. Lanny
; REGISTRATION NUMBER: 27,578
; REFERENCE/DOCKET NUMBER: 57271A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 722-9332
; TELEFAX: (716) 477-4545
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 nucleotides
; TYPE: Nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: Primer for HIV-1 DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: No
; ORIGINAL SOURCE: Synthetically prepared
; IMMEDIATE SOURCE: Same
; PUBLICATION INFORMATION: Unknown
US-08-495-741-25

```

```

Query Match      29.7%; Score 17.8; DB 1; Length 40;
Best Local Similarity 75.9%; Pred.No. 2,3e-03;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      1 GACATCCGCCAGGCGCCGACGAGGAGCCCTT 29
        |||||  |||||  |||||  |||||  |||||
Db      32 GACATAGACAGGAGGACCAAGGAGCCCTT 4

```

Search completed: September 22, 2003, 14:10:34  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Computer Ltd.

OK nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 ; Search time 1380 seconds  
(without alignments)  
1778.679 Million cell updates/sec

Title: US-09-475-704A-2  
Perfect score: 60  
Sequence: 1 gacatccgcagagccca.....tggacggctcttcagacc 60

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapex: 1.0

Searched: 2868711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 1520254

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match: 0%  
Maximum Match: 100%  
Listing first 45 summaries

Database :

GenEmbl :

- 1: gb\_ba.\*
- 2: gb\_hlg.\*
- 3: gb\_in.\*
- 4: gb\_sm.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_or.\*
- 21: em\_ov.\*
- 22: em\_lo.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vl.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	6	AX455886	AX455886 Sequence
2	55.2	92.0	60	6	AX455885	AX455885 Sequence
3	55.2	92.0	60	6	AX455937	AX455937 Sequence
4	34.4	57.3	60	6	AX455910	AX455910 Sequence
5	32.8	54.7	60	6	AX455938	AX455938 Sequence
6	21.9	36.3	84	6	AK021158	AK021158 Sequence
7	21.4	35.7	97	7	AF223423	AF223423 Pseudomon
8	21.2	35.3	84	9	HS091013	U91013 Homo sapien
9	20.8	34.7	84	6	AK021157	AK021157 Sequence
10	20.8	34.7	84	6	AK021159	AK021159 Sequence
11	20.4	34.0	60	6	AR169921	AR169921 Sequence
12	20	33.3	81	6	AX480659	AX480659 Sequence
13	19.6	32.7	96	6	AX088790	AX088790 Sequence
14	19.4	32.3	51	9	HSDNASFAT	Z22386 H.sapiens D
15	19.4	32.3	53	6	I08503	I08503 Sequence 17
16	19.4	32.3	65	6	I08502	I08502 Sequence 16
17	19.4	32.3	78	6	AX480660	AX480660 Sequence
18	19.2	32.0	51	6	AX158488	AX158488 Sequence
19	19.2	32.0	100	9	AJ81116	AJ008116 Homo sapi
20	18.8	31.3	30	6	AX026570	AX026570 Sequence
21	18.8	31.3	74	5	AY227159	AY227159 Alca ford
22	18.6	31.0	51	6	AX157786	AX157786 Sequence
23	18.6	31.0	60	6	AX701057	AX701057 Sequence
24	18.4	30.7	51	6	AX203926	AX203926 Sequence
25	18.4	30.7	62	6	A10075	A10076 Epstein-Bar
26	18.4	30.7	62	14	EBVCLGAS	XC7631 EBV B95-8 C
27	18.4	30.7	69	10	AY205871	AY205871 Mus muscu
28	18.2	30.3	88	6	AX146316	AX146316 Sequence
29	18.2	30.3	95	6	AX523879	AX523879 Sequence
30	18.2	30.3	38	6	AR164537	AR164537 Sequence
31	18	30.0	72	6	AR051746	AR051746 Sequence
32	18	30.0	72	9	HSU91219	U91219 Homo sapien
33	18	30.0	80	14	AF040847	AF040847 Hepatitis
34	18	30.0	80	14	AF040849	AF040849 Hepatitis
35	18	30.0	98	3	ASPBP23A2	AF134278 Agrotis s
36	18	30.0	98	3	ASPBP23B2	AF134280 Agrotis s
37	18	30.0	98	3	ASPBP24A2	AF134282 Agrotis s
38	18	30.0	98	3	ASPBP25A2	AF134284 Agrotis s
39	18	30.0	98	3	ASPBP26A2	AF134286 Agrotis s
40	18	30.0	98	3	ASPBP27A2	AF134288 Agrotis s
41	17.8	29.7	40	6	AR142031	AR142031 Sequence
42	17.8	29.7	40	6	I59898	I59898 Sequence 25
43	17.8	29.7	40	6	I68065	I68065 Sequence 3
44	17.8	29.7	40	6	I96756	I96756 Sequence 25
45	17.8	29.7	40	6	I95781	I95781 Sequence 25

ALIGNMENTS

RESULT 1	AX455886	AX455886	50 bp	DNA	linear	PAT 06-JUL-2002
LOCUS	Sequence 2 from Patent WO0204493.					
DEFINITION	AX455886					
ACCESSION	AX455886					
VERSION	AX455886.1	GI:21714879				
KEYWORDS	Human immunodeficiency virus					
SOURCE	Human immunodeficiency virus					
ORGANISM	Viruses; Retroviridae; Retroviridae; Lentivirus; Primate					
REFERENCE	1					
AUTHORS	zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.					
TITLE	Polynucleotides encoding antigenic hiv type c polypeptides, polypeptides and uses thereof					



```

Query Match: 54.7%; Score 32.8; DB 6; Length: 60;
Best Local Similarity 71.7%; Pred. No. 45;
Matches 43; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GACATCGCGCCAGGCCCAAGGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAGACC 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 GACATAAACCAAGGCCCAAGGAGCCCTTGTAGAGACTGTAGACCGGTTCTTAAAC 60

RESULT 6
AR021158 LOCUS 84 bp DNA linear PAT 05-DEC-1999
DEFINITION Sequence 7 from patent US 5789389.
ACCESSION AR021158
VERSION AR021158.1 GI:3975773
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 84)
AUTHORS Tarasiewicz,D.G., Schott,B., Holzmayer,T.A. and Robinson,I.P.
TITLE BC12 derived genetic elements associated with sensitivity to
chemotherapeutic drugs
JOURNAL Patent: US 5789389-A 7 04-AUG-1998;
FEATURES
    source
        1..84
            /organism="unknown"
BASE COUNT 11 a 38 c 23 g 12 t
ORIGIN

Query Match: 36.3%; Score 21.8; DB 6; Length 84;
Best Local Similarity 61.4%; Pred. No. 5.5e+04;
Matches 35; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 3 CATCGCGAGGCCCAAGGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAGAC 59
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 15 COTCGCGAGCGCGGACGACTTCTCGCGCGGCTACGCGCGGAC:TCGCGGAGAC 7;

RESULT 7
AF223423 LOCUS 97 bp DNA linear BCT 29-MAR-2000
DEFINITION Pseudomonas aeruginosa clone 570 biofilm-dependent regulatory
element.
ACCESSION AF223423
VERSION AF223423.1 GI:7540377
KEYWORDS
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE 1 (bases 1 to 97)
AUTHORS Meyers,J.B., van Zyl,W.H. and Brozel,V.S.
TITLE Pseudomonas aeruginosa contains a novel and diverse group of
biofilm-dependent regulatory elements
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 97)
AUTHORS Meyers,J.B. and Brozel,V.S.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-2000) Microbiology and Plant Pathology,
University of Pretoria, Lunnon Road, Pretoria, Gauteng 0002, South
Africa
FEATURES
    source
        1..97
            /organism="Pseudomonas aeruginosa"
            /mol_type="genomic DNA"
            /strain="PAOI"
            /db_xref="taxon:287"
            /clone="570"
        misc_feature
            1..97
                /note="biofilm-dependent regulatory element; up-regulated
                when grown as a biofilm on a solid surface."
BASE COUNT 16 a 35 c 35 g 11 t

```

## ORIGIN

```

Query Match: 35.7%; Score 21.4; DB 1; Length 97;
Best Local Similarity 61.8%; Pred. No. 7e-04;
Matches 34; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GACATCGCGCAGGCCCAAGGAGCCCTTCGGGACTACGTGGACCGCTTCTTCA 55
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 17 GUCCACCGCGCGGCATCGAGGTGCGGTCGACGAGCGCGAGCGCGCTTCTCA 71

RESULT 8
HSU91013 LOCUS 84 bp mRNA linear PRI 15-JUL-1997
DEFINITION Homo sapiens clone FEL399 T-cell receptor delta chain (TCRDVLJ3)
mRNA, partial cds.
ACCESSION U91013
VERSION U91013.1 GI:2239433
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 84)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Holtmeier,W., Witthoft,T., Hennemann,A., Winter,H.S. and
Kagnoff,M.F.
TITLE The TCR-delta repertoire in human intestine undergoes
characteristic changes during fetal to adult development
JOURNAL J. Immunol. 158 (12), 5632-5641 (1997)
MEDLINE 57334214
PUBMED 9190911
REFERENCE 2 (bases 1 to 84)
AUTHORS Holtmeier,W., Witthoft,T., Hennemann,A., Harland,S.W. and
Kagnoff,M.F.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-1997) Department of Medicine, University of
Frankfurt, Theodor-Stern Kai #7, Frankfurt 60590, Germany
FEATURES
    source
        1..84
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /chromosome="14"
            /clone="FEL399"
            /tissue_type="small intestine"
            /dev_stage="fetus; 20 weeks gestation"
            <1..>84
            /gene="TCRDVLJ3"
            <1..>84
            /gene="TCRDVLJ3"
            /note="rearranged; contains CDR3 domain, 16 amino acids
            (calculation according to: Rock,E., J.Exp.Med.,
            179:323-328,1994)"
            /codon_start=1
            /product="T-cell receptor delta chain"
            /protein_id="AAC51576.1"
            /db_xref="GI:2239434"
            /translation="LEDKAKYFCALGEPTYWIGSGWDTQMF"
            <1..>84
            /gene="TCRDVLJ3"
BASE COUNT 20 a 18 c 25 g 21 t
ORIGIN

Query Match: 35.3%; Score 21.2; DB 9; Length 84;
Best Local Similarity 69.0%; Pred. No. 8.1e+04;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 ACATCGCGCAGGCCCAAGGAGCCCTTCGGGACTACGTGG 43
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 82 ACATCTGTCGGGTGTCCAGGAGCCTATCCCCAGTACGTGG 41

RESULT 9

```

```

AR021157
LOCUS       AR021157               84 bp    DNA             linear      PAT 05-SEP-1998
DEFINITION   Sequence 5 from patent US 5789389.
ACCESSION   AR021157
VERSION     AR021157.1  GI:3975772
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 84)
AUTHORS    Tarasiewicz,D.G., Schott,B., Holzmayr,T.A. and Robinson,L.B.
TITLE      BCL2 derived genetic elements associated with sensitivity to
           chemotherapeutic drugs
JOURNAL     Patent: US 5789389-A 5 04-AUG-1998;
FEATURES    Location/Qualifiers
             ..84
             /organism="unknown"
BASE COUNT  11 a 37 c 23 g 13 t
ORIGIN
Query Match      34.7%; Score 20.8; DB 6; Length 84;
Best Local Similarity 60.7%; Pred. No. 1.1e+05;
Matches 34; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY  3  CATCCGCCAGGCGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGA 58
    | ||||| ||| | || |||| ||| | |||| ||| | |||| ||| |
Db  15  CTTCCGCCAGGCGCGGAGCACTTCCTCCGCGCTACGCGCGGACTTCGCGGAGA 70

RESULT 10
LOCUS       AR021159               84 bp    DNA             linear      PAT 05-DEC-1998
DEFINITION   Sequence 9 from patent US 5789389.
ACCESSION   AR021159
VERSION     AR021159.1  GI:3975774
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 84)
AUTHORS    Tarasiewicz,D.G., Schott,B., Holzmayr,T.A. and Robinson,L.B.
TITLE      BCL2 derived genetic elements associated with sensitivity to
           chemotherapeutic drugs
JOURNAL     Patent: US 5789389-A 9 04-AUG-1998;
FEATURES    Location/Qualifiers
             1..84
             /organism="unknown"
BASE COUNT  10 a 37 c 23 g 14 t
ORIGIN
Query Match      34.7%; Score 20.6; DB 6; Length 84;
Best Local Similarity 60.7%; Pred. No. 1.1e+05;
Matches 34; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY  3  CATCCGCCAGGCGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGA 58
    | ||||| ||| | || |||| ||| | |||| ||| | |||| ||| |
Db  15  CTTCCGCCAGGCGCGGAGCACTTCCTCCGCGCTACGCGCGGACTTCGCGGAGA 70

RESULT 11
LOCUS       AR169921               60 bp    DNA             linear      PAT 17-DEC-2001
DEFINITION   Sequence 7 from patent US 6291208.
ACCESSION   AR169921
VERSION     AR169921.1  GI:17907380
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 60)
AUTHORS    Anand,N.N., Barber,B.H., Cates,G.C., Caterini,J.E. and Klein,M.H.
TITLE      Chimeric antibodies for delivery of antigens to selected cells of
           the immune system

```

```

JOURNAL     Patent: US 6291208-A 7 18-SEP-2001;
FEATURES    Location/Qualifiers
             1..60
             /organism="unknown"
BASE COUNT  19 a 8 c 12 g 21 t
ORIGIN
Query Match      34.0%; Score 20.4; DB 6; Length 60;
Best Local Similarity 65.2%; Pred. No. 1.4e+05;
Matches 30; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY  13  GGGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGA 58
    | ||||| ||| | || |||| ||| | |||| ||| | |||| ||| |
Db  13  GGTCTTAAGAACCCTTTAGAGACTATCTTGATAGGTTTATAAGA 58

RESULT 12
LOCUS       AX480659               81 bp    DNA             linear      PAT 12-AUG-2002
DEFINITION   Sequence 47 from Patent WO0248189.
ACCESSION   AX480659
VERSION     AX480659.1  GI:22217408
KEYWORDS    .
SOURCE      synthetic construct
           synthetic construct
           artificial sequences.
REFERENCE   1
AUTHORS    Etzerodt,M., Holtet,T.L., Graversen,N.J. and th Gersen,H.C.
TITLE      Combinatorial libraries of proteins having the scaffold structure
           of c-type lectin-like domains
JOURNAL     Patent: WO 0248189-A 47 20-JUN-2002;
           Borean Pharma A/S (DK)
FEATURES    Location/Qualifiers
             1..81
             /organism="synthetic construct"
             /mol_type="genomic DNA"
             /db_xref="taxon:32630"
             /note="oligonucleotide"
BASE COUNT  12 a 20 c 22 g 9 t 18 others
ORIGIN
Query Match      33.3%; Score 20; DB 6; Length 81;
Best Local Similarity 44.2%; Pred. No. 1.8e+05;
Matches 23; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

QY  1  GACATCCGCCAGGCGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCT 52
    | ||||| ||| | || |||| ||| | |||| ||| | |||| ||| |
Db  2  GAGATCTGGTGGCCCTCACGACNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 53

RESULT 13
LOCUS       AX088790/c              56 bp    DNA             linear      PAT 17-MAR-2001
DEFINITION   Sequence 116 from Patent WO0114416.
ACCESSION   AX088790
VERSION     AX088790.1  GI:13397586
KEYWORDS    .
SOURCE      synthetic construct
           synthetic construct
           artificial sequences.
REFERENCE   1
AUTHORS    Neepser,M.P., McClements,W.L., Jansen,K.U., Schultz,L.D., Chen,L.
           and Wang,X.M.
TITLE      Synthetic human papillomavirus genes
JOURNAL     Patent: WO 0114416-A 116 01-MAR-2001;
           Merck & Co., Inc. (US)
FEATURES    Location/Qualifiers
             1..96
             /organism="synthetic construct"
             /mol_type="genomic DNA"
             /db_xref="taxon:32630"
             /note="Codon-Optimized HPV6 E2 fragment"
BASE COUNT  13 a 23 c 40 g 20 t

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 : Search time 193 Seconds

(without alignments)  
839,204 Million cell updates/sec

Title: US-09-475-704A-2

Perfect score: 60

Sequence: 1 qacatccgacggccca.....tgaccqcttcttcaagacc 40

Scoring table: IDENTITY\_NJC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 2722628

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_13Jun01.\*

	1:	2:	3:	4:	5:	6:	7:	8:	9:	10:	11:	12:	13:	14:	15:	16:	17:	18:	19:	20:	21:	22:	23:	24:	25:	
1:	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1980.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1981.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1982.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1983.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1984.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1985.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1986.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1987.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1988.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1989.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1990.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1991.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1992.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1993.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1994.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1995.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1996.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1997.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1998.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA1999.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA2000.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA2001.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA2002.DAT.*	/SIDSL/qcgdata/geneseq/geneseq-emb1/NA2003.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	60	100.0	60	21	AAA51608 HIV Gag major homo
2	58.4	97.3	60	21	AAA70426 HIV p55 Gag Major
3	55.2	92.0	60	21	AAA51607 HIV Gag major homo
4	55.2	92.0	60	24	ABL39974 Synthetic Gag poly
5	34.4	57.3	60	21	AAA51627 Major homology reg
6	33	55.0	90	24	ABK36257 HIV DNA encoding G
7	32.8	54.7	60	24	ABL39975 Wild type 8_5_TVL
8	31.8	53.0	87	7	AAAN60663 HIV virus gag reqi

3	29.8	49.7	90	24	ABK36256 HIV DNA encoding G
10	21.8	36.3	84	17	AAI33692 Nucleic acid encod
11	21	35.0	87	22	ABA71782 Human foetal liver
12	21	35.0	87	22	AAK20151 Human brain expres
13	21	35.0	87	22	AAK46211 Human bone marrow
14	21	35.0	87	22	AAI52117 Probe #20803 used
15	21	35.0	87	23	ABS45930 Human liver sing,e
16	21	35.0	87	24	ABS20522 Human genome-deriv
17	20.8	34.7	84	17	AAI33691 Nucleic acid encod
18	20.8	34.7	84	17	AAI33693 Nucleic acid encod
19	20.4	34.0	60	24	AAI32141 CTLB36.1 antigen p
20	20.4	34.0	65	24	ABN52351 Mouse spliced tran
21	20.2	33.7	65	24	ABN57741 Mouse spliced tran
22	20	33.3	81	24	ABS66623 Loop 1b oligonucle
23	19.6	32.7	96	22	AAF75498 Codon-optimised HP
24	19.4	32.3	45	22	AAI09440 HIV reverse trans
25	19.4	32.3	53	18	AAI71251 Basic fibroblast g
26	19.4	32.3	65	18	AAI71250 Basic fibroblast g
27	19.4	32.3	78	24	ABS66624 Loop 1c oligonucle
28	19.4	32.3	100	24	ABS69273 Novel murine polyn
29	19.2	32.0	51	22	AAI74875 Human silent SNP c
30	18.8	31.3	30	21	AAA46684 Primer R2 used to
31	18.8	31.3	51	21	AAA76375 Human calmodulin-l
32	18.6	31.0	51	22	AAI74173 Human silent SNP c
33	18.6	31.0	60	24	ABN36996 Human spliced tran
34	18.4	30.7	37	13	AAQ21583 Sequence of primer
35	18.4	30.7	41	24	ABA04315 Human zinc finger
36	18.4	30.7	50	22	AAI34430 Human SNP oligonuc
37	18.4	30.7	51	22	AAI32838 Human SNP containi
38	18.4	30.7	51	22	AAH79417 Human DNA containi
39	18.4	30.7	60	24	ABN40591 Human spliced tran
40	18.4	30.7	62	10	AAH91755 DNA sequence which
41	18.2	30.3	52	24	ABK86871 Topoisomerase vect
42	18.2	30.3	60	24	ABN35235 Human spliced tran
43	18.2	30.3	89	24	AAI44429 West Nile virus ca
44	18.2	30.3	99	24	ABS68080 B-domain-deleted-F
45	18	30.0	38	20	AAZ25414 Infectious pancrea

#### ALIGNMENTS

RESULT 1  
AAA51608 standard; DNA; 60 BP.  
XX AAA51608;  
AC AAA51608;

31-OCT-2000 (first entry)

HIV Gag major homology region nucleotides 841-900.

Gag: expression cassette; antigenic; type C; HIV: Env; synthetic;  
DNA immunization; packaging cell line; antigen presentation; ss.

Human immunodeficiency virus.

WC200039304-A2.

06-JUL-2000.

30-DEC-1999; 99WO-US31273.

31-DEC-1998; 98US-0114495.

01-SEP-1999; 99US-0152195.

(CHIR ) CHIRON CORP.

Barnett S, Zur Megede J;

WPI; 2000-452401/39.

Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV

PT Env polypeptide and the polypeptide useful for immunizing a mammal:  
 especially human against HIV

XX  
 XX Claim 1; Page 92; 113pp; English.

XX Expression cassettes comprising a polynucleotide encoding antigenic  
 CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are  
 CC used in DNA immunization, generation of packaging cell lines and  
 CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag  
 CC expression cassettes exhibit increased potency for induction of  
 CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1  
 CC self-assemble into non-infectious virus-like particles which are used as  
 CC a matrix for the proper presentation of an antigen enveloped or  
 CC associated to the immune system of the host.

XX Sequence 60 bp; 12 A; 24 C; 15 G; 9 T; 0 other;

SQ Query Match 100.0%; Score 60; DB 21; Length 60;  
 Best local Similarity 100.0%; Pred. No. 3e+09;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATCGCCGAGGGGCCCAAGGAGGCCTTCCGGACTAGTGAGCGGTTCAGACAC 60  
 Db 1 GACATCGCCGAGGGGCCCAAGGAGGCCTTCCGGACTAGTGAGCGGTTCAGACAC 60

RESULT 2  
 AAA70426

ID AAA70426 standard; DNA; 60 BP.  
 AC  
 XX AAA70426;  
 AC

DT 28-NOV-2000 (first entry)  
 XX  
 DT HIV p55 Gag Major Homology Region coding sequence.  
 DE  
 DE HIV-1; AIDS; Gag; vaccine; expression cassette; ss.  
 KW  
 XX Human immunodeficiency virus type 1.  
 OS

XX Key location/Qualifiers  
 FT 1..60  
 CDS /tag= a  
 FT /partial  
 FT /product= "HIV p55 Gag Major Homology Region protein"  
 FT /note= "No stop codon given"

XX WO200039302-A2.  
 PN  
 XX  
 XX 06-JUL-2000.  
 PD  
 XX 30-DEC-1999; 99WO-US31245.  
 PF  
 XX 31-DEC-1999; 98CS-0114495.  
 PR  
 XX 01-DEC-1999; 99CS-016847.  
 XX  
 XX (CHIR ) CHIRON CORP.  
 PA  
 XX Barnett S., Zur Megede J., Srivastava I., Lian Y., Hartog K., Liu H.;  
 PI Greer C., Selby M., Walker C;  
 XX  
 XX WPI: 2000-452400/39.  
 DR P-PSDB; AAB14214.  
 DR  
 XX Expression cassettes encoding the human immunodeficiency virus (HIV)  
 PT Gag-containing polypeptide useful for vaccinating against HIV  
 PT infections and acquired immunodeficiency syndrome (AIDS).  
 XX  
 XX Claim 1; Page 345; 391pp; English.  
 PS  
 XX The present invention relates to synthetic HIV Gag expression cassettes.  
 CC The Gag protein of HIV is needed for the assembly of virus-like  
 CC particles. In addition, the gag protein is involved in many stages of the

```

CC HIV life cycle, including assembly, virion maturation after particle
CC release and early post-entry steps in viral replication. The expression
CC cassette may be used for the recombinant expression of HIV
CC Gag-polypeptides which may then be used to vaccinate against HIV
CC infection and acquired immunodeficiency syndrome (AIDS). The present
CC sequence is the coding sequence of HIV p55 Gag Major Homology Region.
CC This sequence is located within the p24-CA sequence of Gag. Mutations in
CC this region can severely impair particle production.
XX
XX Sequence 60 BP: 13 A; 24 C; 15 G; 8 T; 0 other;
XX
Query Match          97.3%; Score 58.4; DB 21; Length 60;
Best Local Similarity 98.3%; Pred No. 9.3e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
2Y : GACATCGCCGACGGGCTCCAGAGAGCCCTTCGCGACTACG:GGAGCGCTTCTTCAAGACC 60;
DB : |||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
   : GACATCGCCGACGGGCTCCAGAGAGCCCTTCGCGACTACG:GGAGCGCTTCTTCAAGACC 60;
XX
RESULT 3
AAAS1607
ID AAAS1607 standard; DNA: 60 BP.
XX
XX AAAS1607:
XX
XX 31-DEC-2000 (first entry)
XX
XX HIV Gag major homology region nucleotides 844-903.
XX
XX Gag; expression cassette; antigenic; type C; HIV; Env; Synthetic;
XX DNA immunization; packaging cell line; antigen presentation; ss.
XX
XX Human immunodeficiency virus.
XX
XX WC2000C9304-A2.
XX
XX 06-JUL-2000.
XX
XX 30-DEC-1999; 99WO-US31273.
XX
XX 31-DEC-1998; 98US-0114495.
XX 01-SEP-1999; 99US-0152195.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Barnett S. Zur Megede J.
XX
XX WPI; 2030-452401/39.
XX
XX Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
XX Env polypeptide and the polypeptide useful for immunizing a mammal
XX especially human against HIV
XX
XX C-aim 1; Page 92; 113pp; English.
XX
XX Expression cassettes comprising a polynucleotide encoding antigenic
XX type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
XX useful in DNA immunization, generation of packaging cell lines and
XX production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
XX expression cassettes exhibit increased potency for induction of
XX cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
XX self-assemble into non-infectious virus-like particles which are used as
XX a matrix for the proper presentation of an antigen entrapped or
XX associated to the immune system of the host.
XX
XX Sequence 60 BP: 14 A; 22 C; 15 G; 9 T; 0 other;
XX
Query Match          92.0%; Score 55.2; DB 21; Length 60;
Best Local Similarity 95.0%; Pred. No. 8.7e-08;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
1 GACATCGCCGACGGGCTCCAGAGAGCCCTTCGCGACTACG:GGAGCGCTTCTTCAAGACC 60;

```

```

Db      1 GACATCAAGCAGGCCCCAAGAGCCCTTCCCGACTAGCTGGACCGCTTCTTCAAGACC 60
        |||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
RESULT 4
ABL39974
ID      ABL39974 standard; DNA; 60 BP.
XX
AC      ABL39974;
XX
DT      15-MAY-2002 (first entry)
XX
DE      Synthetic Gag polynucleotide sequence SEQ ID NO:53.
XX
KW      Human immunodeficiency virus type C; antigenic HIV type C protein;
KW      immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW      immunostimulant; gene therapy; gene; ds.
XX
XX      Human immunodeficiency virus type C.
OS      Synthetic.
XX
XX      WO200204453-A2.
PN
XX
PD      17-JAN-2002.
XX
PF      05-JUL-2001; 2001WO-US21241.
XX
PR      05-JUL-2000; 2000US-0610313.
XX
PA      (CHIR ) CHIRON CORP.
XX
PA      (UYST-) UNIV STELLENBOSCH.
XX
PI      Zur Megede J, Barnett SW, Engebrecht S, Van Rensburg EJ;
XX
XX      WPI; 2002-154920/20.
DR
XX
PT      New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT      in applications including DNA immunization or generation of packaging
PT      cell lines, particularly in gene therapy.
XX
PS      Example 1; Fig 24; 233pp; English.
XX
XX      The present invention describes expression cassettes comprising a
XX      polynucleotide sequence encoding a polypeptide comprising immunogenic
XX      HIV type C polypeptides. The expression cassettes comprise any of the
XX      HIV type C sequences encoding gag, pol, vif, vpr, tat, rev, vpu, env or
XX      nef (1). (1) have immunostimulant activity and can be used in gene
XX      therapy. The HIV type C polynucleotides are useful in applications
XX      including DNA immunisation, generation of packaging cell lines, and
XX      production of HIV type C proteins. The polynucleotides are particularly
XX      useful in gene therapy and DNA immunisation applications. ABL39942 to
XX      ABL40054 and ABR06204 to ABR06215 represent sequences used in the
XX      exemplification of the present invention.
XX
SQ      Sequence 60 BP; 14 A; 22 C; 15 G; 9 T; 0 other;
        Query Match          92.0%; Score 55.2; DB 24; Length 60;
        Best Local Similarity 95.0%; Pred. No. 6.7e-08;
        Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GACATCCGCGAGGCCCAAGAGCCCTTCCCGACTAGCTGGACCGCTTCTTCAAGACC 60
        |||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db      1 GACATCAAGCAGGCCCCAAGAGCCCTTCCCGACTAGCTGGACCGCTTCTTCAAGACC 60
        |||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
RESULT 5
AAA51627
ID      AAA51627 standard; DNA; 60 BP.
XX
XX      AAA51627;
AC
XX
XX      31-OCT-2000 (first entry)
DT
XX

```

```

DE      Major homology region of HIV-1 p55 (gag) coding sequence.
XX
XX      gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW      DNA immunization; packaging cell line; antigen presentation; ss.
XX
XX      Human immunodeficiency virus type C strain AF110965.
OS
XX      WO200039304-A2.
PN
XX
PD      06-JUL-2000.
XX
XX      30-DEC-1999; 99WO-US31273.
XX
XX      31-DEC-1998; 98US-0114495.
PR
XX      01-SEP-1999; 99US-0152195.
XX
XX      (CHIR ) CHIRON CORP.
PA
XX      Barnett S, Zur Megede J;
PI
XX      WPI; 2000-452401/39.
DR
XX      P-PSDB; AAY96947.
XX
XX      Polynucleotide encoding antigenic type C HIV gag polypeptide or a HIV
XX      Env polypeptide and the polypeptide useful for immunizing a mammal
XX      especially human against HIV
XX
XX      Example 1; Page 112; 113pp; English.
XX
XX      Expression cassettes comprising a polynucleotide encoding antigenic
XX      type C human immunodeficiency virus (HIV) gag or Env polypeptides are
XX      useful in DNA immunization, generation of packaging cell lines and
XX      production of gag- and/or Env-containing proteins. Synthetic Env and gag
XX      expression cassettes exhibit increased potency for induction of
XX      cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
XX      self-assemble into non-infectious virus-like particles which are used as
XX      a matrix for the proper presentation of an antigen entrapped or
XX      associated to the immune system of the host.
XX
SQ      Sequence 60 BP; 24 A; 13 C; 11 G; 12 T; 0 other;
        Query Match          57.3%; Score 34.4; DB 21; Length 60;
        Best Local Similarity 73.3%; Pred. No. 0.18;
        Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY      1 GACATCCGCGAGGCCCAAGAGCCCTTCCCGACTAGCTGGACCGCTTCTTCAAGACC 60
        |||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db      1 GACATAAACAAGGACCAAAAGAGCCCTTTAGAGACTATGAGACCGGCTTCTTAAACC 60
        |||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
RESULT 6
ABK36257
ID      ABK36257 standard; DNA; 90 BP.
XX
XX      ABK36257;
AC
XX
XX      08-MAY-2002 (first entry)
DT
XX
XX      HIV DNA encoding gag segment 20.
DE
XX
XX      Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW      viral infection; human immunodeficiency virus; melanoma;
KW      bacterial infection; Salmonella; Legionella; parasitic infection;
KW      Trypanosoma; Toxoplasma; Giardia; ds.
XX
XX      Human immunodeficiency virus type 1.
OS      Synthetic.
XX
XX      WO200190197-A1.
PN
XX
XX      29-NOV-2001.
PD
XX
XX      25-MAY-2001; 2001WO-AU00622.
PF

```

XX  
PR 26-MAY-2000: 2000AU-0007761.  
XX

PA (AUSU ) UNIV AUSTRALIAN NAT.  
XX

PI Thomson SA, Ramshaw IA;  
XX

XX WPI: 2002-147575/19.  
XX

DR P-PSDB; AAU84418.  
XX

XX New synthetic polypeptides having several different segments of at  
PI least one parent polypeptide linked together differently compared to  
FT the linkage in the parent polypeptide, for inducing immune response  
PT against a pathogen or cancer  
XX

PS Example 1: Fig 12: 364pp; English.  
XX

XX The invention relates to a new synthetic polypeptide (I) comprising  
CC several different segments of at least one parent polypeptide linked  
CC together in a different relationship relative to their linkage in the  
CC parent polypeptide to impede, abrogate or otherwise alter at least one  
CC function associated with the parent polypeptide and for inducing an  
CC immune response against a pathogen or cancer. Also included are a  
CC synthetic polynucleotide encoding and a computer system for  
CC designing the synthetic polypeptides. The synthetic polypeptides and  
CC polynucleotides are referred to as a vaccine. The synthetic polypeptide is  
CC useful for modulating immune responses preferably directed against a  
CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
CC liver, esophagus, brain, testicle, uterus), as potentiating agents.  
CC compositions comprising the polypeptide may be used in the treatment or  
CC prophylaxis against viral (such as infections caused by HIV (human  
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacteria:  
CC (e.g., infections caused by Neisseria, Meningococcus, Haemophilus,  
CC Salmonella, Streptococcus, Legionella and Mycobacterium) or parasitic  
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
CC sequence encodes a peptide derived from a parent protein used to  
CC construct a vaccine of the invention.  
XX

SQ Sequence 90 BP; 30 A; 20 C; 22 G; 14 T; 4 other;  
XX

Query Match 55.0%; Score 33; DB 24; Length 90;  
Best Local Similarity 76.5%; Pred. No. 0.48;  
Matches 36; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 10 CAGGGCCCCAAGAGAGCCCTTCGCGACTAGTGGACGGCTTCTCAAGACC 60  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 4 CAGGGACCCAAAGAGAGCCCTTCAGAGACATGATGATAGGTTTWCANACC 54  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7

ABL39975  
ID ABL39975 standard; DNA: 60 BP.  
XX

AC ABL39975;  
XX

XX 15-MAY-2002 (first entry)  
XX

XX Wild type 8\_5\_TV1\_C-2A Gag major homology region DNA SEQ ID NO:54.  
XX

XX Human immunodeficiency virus type C; antigenic HIV type C protein;  
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;  
KW immunostimulant; gene therapy; gene; ds.  
XX

OS Human immunodeficiency virus type C.  
XX

XX WO200204493-A2.  
XX

XX 17-JAN-2002.  
XX

XX 05-JUL-2001; 2001WO-US21241.  
XX

XX 05-JUL-2000: 2000US-0610313.  
XX

PA (CHIR ) CHIRON CORP.  
XX

PA (JYST-) UNIV STELLENBOSCH.  
XX

XX Zur Megele J, Barnett SW, Engelbrecht S, Van Rensburg EJ;  
XX

XX WPI: 2002-154920/20.  
XX

XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful  
PI in applications including DNA immunization or generation of packaging  
FT cell lines, particularly in gene therapy  
PT

PS Example 1: Fig 25: 233pp; English.  
XX

XX The present invention describes expression cassettes comprising a  
CC polynucleotide sequence encoding a polypeptide comprising immunogenic  
CC HIV type C polypeptides. The expression cassettes comprise any of the  
CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or  
CC Nef (-). (i) have immunostimulant activity and can be used in gene  
CC therapy. The HIV type C polynucleotides are useful in applications  
CC including DNA immunisation, generation of packaging cell lines, and  
CC production of HIV type C proteins. The polynucleotides are particularly  
CC useful in gene therapy and DNA immunisation applications. ABL39942 to  
CC ABL40054 and ABL06204 to ABL06215 represent sequences used in the  
CC exemplification of the present invention.  
XX

SQ Sequence 60 BP; 24 A; 13 C; 11 G; 12 T; 0 other;  
XX

Query Match 54.7%; Score 32.8; DB 24; Length 60;  
Best Local Similarity 71.7%; Pred. No. 0.54;  
Matches 43; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GACATCCGCGCCAGGCGCCCAAGAGAGCCCTTCGCGACTAGTGGACGGCTTCTCAAGACC 60  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1 GACATAAACAAAGAGCGCCAAAGAACCCCTTAGAGACTATGTAGACGGGTCTTTAAACC 60  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8

AAN50663  
ID AAN50663 standard; DNA: 87 BP.  
XX

AC AAN50663;  
XX

XX 25-MAR-2003 (updated)  
XX

XX 03-OCT-2002 (updated)  
XX

XX 12-JUN-1991 (first entry)  
XX

XX HIV virus gag region encoding part of the p25 protein.  
XX

XX HIV virus; AIDS; vaccine; ss.  
XX

XX Human immunodeficiency virus.  
XX

XX Key Location/Qualifiers  
XX

XX mat\_peptide i..87  
XX

XX /\*tag=p25 protein region of the HIV virus  
XX

XX WC8606414-A.  
XX

XX 06-NOV-1986.  
XX

XX 21-APR-1986; 86WO-US00831.  
XX

XX 26-MAR-1986; 86US-0844485.  
XX

XX 29-APR-1985; 85US-0728052.  
XX

XX 19-AUG-1985; 85US-0767303.  
XX

XX (GENE-) GENETIC SYSTEMS CORP.  
XX

XX Cosand WL;  
XX





Db 86 AGGAGCCCATCCACACTACCTGACCATCATGATCAA 50  
||||| ||| |||| |||| |||| ||||

## RESULT 13

AAK46211/c  
ID AAK46211 standard; DNA: 57 BP.

XX AAK46211;

XX 06-NOV-2001 (first entry)

XX Human bone marrow expressed single exon probe SEQ ID NO: 20768.

XX Human bone marrow expressed exon; gene expression analysis: probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

OS Homo sapiens.

PN WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-C207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow.

XX Example 4; SEQ ID NO: 20768; 658pp - Sequence Listing; English.

PS The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukaemia and myeloma. The present sequence is one of

CC the probes of the invention.

XX Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;

## RESULT 14

AAI52117/c

ID AAI52117 standard; DNA: 87 BP.

XX AAI52117;

XX 17-OCT-2001 (first entry)

XX Probe #20803 used to measure gene expression in human placenta sample.

XX Probe: microarray; human; placenta; antenatal diagnosis;

KW genetic disorder; ss.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-48897/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human placenta.

XX Claim 25; SEQ ID NO 20803; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP).

XX The present sequence is one such probe. The probes are useful for

XX producing a microarray for predicting, measuring and displaying gene

XX expression in samples derived from human placenta. The probes are useful

XX for antenatal diagnosis of human genetic disorders.

XX Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;

XX Query Match 35.0%; Score 21; DB 22; Length 87;

XX Best Local Similarity 73.0%; Pred. No. 2.1e+03;

XX Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

XX 20 AGGAGCCCTTCGGGACTACCTGGACCGCTTCTTCAA 56

||||||| ||| ||||| ||||| |||||

86 AGGAGCCCATCCACACTACCTGACCATCATGATCAA 50

ARS45930/c

ID ARS45930 standard; DNA: 87 BP.

XX ARS45930;

XX 25-FEB-2003 (first entry)

XX Human liver single exon probe, SEQ ID NO 20920.

XX Human; single exon nucleic acid probe; liver; cirrhosis;

XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;

XX coronary heart disease; ss.

XX Homo sapiens.

XX WO200157273-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00664.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

```

PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analysing gene expression in human adult liver -
XX
XX Claim 4: SEQ ID No 20920; 658bp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (1) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 1309 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (1) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ARS25011-ABSS1005 represent
CC human liver single exon nucleic acid probes of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;

```

```

Query Match          35.0%; Score 21; DB 23; Length 87;
Best Local Similarity 73.0%; Pred. No. 2.1e+03;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Caps 0;
QY 20 AGGAGCCCTTCGCGACTACCTGACACGCTCTTCA 56
   ||||| ||| |||| | ||| | |||
Db 86 AGGAGCCCATCCACACTACCTGAACCATGATCAA 50

```

Search completed: September 22, 2003, 14:08:51  
Job time : 195 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 : Search time 1798.5 seconds  
(without alignments)  
810.825 Million cell updates/sec

Title: US-09-475-704A-2

Perfect score: 60

Sequence: 1 gacatccgcagggccccc.....tgaccgctcttcaagacc 60

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 452990

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba:\*\*

2: em\_esthum:\*\*

3: em\_estin:\*\*

4: em\_estmu:\*\*

5: em\_estov:\*\*

6: em\_estpl:\*\*

7: em\_estro:\*\*

8: em\_estro:\*\*

9: gb\_est1:\*\*

10: gb\_est2:\*\*

11: gb\_est3:\*\*

12: gb\_est3:\*\*

13: gb\_est4:\*\*

14: gb\_est5:\*\*

15: em\_estom:\*\*

16: em\_estom:\*\*

17: em\_gss\_hum:\*\*

18: em\_gss\_inv:\*\*

19: em\_gss\_pln:\*\*

20: em\_gss\_vrt:\*\*

21: em\_gss\_fun:\*\*

22: em\_gss\_mam:\*\*

23: em\_gss\_mus:\*\*

24: em\_gss\_pro:\*\*

25: em\_gss\_rnd:\*\*

26: em\_gss\_rnd:\*\*

27: em\_gss\_vrt:\*\*

28: gb\_gss1:\*\*

29: gb\_gss2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22.2	37.0	88	BH413628	BH413628 1007033FC
2	22.2	37.0	88	BH413629	BH413629 1007033FC
3	21.8	36.3	98	AF039817	AF039817 AF039817
4	21.4	35.7	50	AU108059	AU108059 AU108059

5	21.4	35.7	84	14	H27186	H27186 yll15409.r1
6	21.4	35.7	93	14	W35927	W35927 mc32a07.r1
7	20.8	34.7	92	28	A2920060	A2920060 1006017H0
8	20.4	34.0	58	9	AI048494	AI048494 uc6.a10.r
9	20.4	34.0	95	13	BQ760009	BQ760009 EBP107.SQ
10	20.2	33.7	74	9	AI272982	AI272982 qv63c07.x
11	20.2	33.7	77	9	AW022596	AW022596 df41b06.y
12	20.2	33.7	85	14	R12059	R12059 y54b11.r1
13	20.2	33.7	91	14	T48611	T48611 yb01a10.r1
14	20	33.3	75	29	CC043452	CC043452 3591_1_15
15	20	33.3	79	9	AA772825	AA772825 ae74e12.s
16	19.6	32.7	79	9	AA456718	AA456718 aa13h07.r
17	19.6	32.7	98	14	H49859	H49859 vo24e10.s1
18	19.4	32.3	50	9	AU102742	AU102742 AU102742
19	19.4	32.3	84	9	AA02311C	AA02311C mh6sc10.r
20	19.2	32.0	97	9	AA124117	AA124117 mp91h01.r
21	19	31.7	70	9	AI036945	AI036945 uh23b11.r
22	19	31.7	82	9	AA154612	AA154612 ud33d05.r
23	19	31.7	85	9	AA615572	AA615572 vo79d07.r
24	19	31.7	85	9	AA241182	AA241182 mv27e07.r
25	19	31.7	95	13	B0652348	B0652348 1112039D0
26	18.8	31.3	50	9	AU102354	AU102354 AU102354
27	18.8	31.3	55	12	BM285363	BM285363 EST00004
28	18.8	31.3	55	12	BM493218	BM493218 EST00001
29	18.8	31.3	55	13	BQ094077	BQ094077 045802.36
30	18.6	31.0	43	28	A2417237	A2417237 1M0192M18
31	18.6	31.0	74	29	CC249304	CC249304 RRF072.Ra
32	18.6	31.0	78	29	CC027848	CC027848 3591_1_7_
33	18.6	31.0	63	10	BE139376	BE139376 xr69q02.x
34	18.6	31.0	84	28	BH222409	BH222409 1006107B0
35	18.6	31.0	86	9	AA627888	AA627888 nq70a06.s
36	18.6	31.0	91	12	BG953271	BG953271 602828373
37	18.6	31.0	92	14	F30532	F30532 HSPD21049.H
38	18.6	31.0	95	29	CNS031GR	A245556 Tetraodon
39	18.6	31.0	100	29	CC027859	CC027859 3591_1_7_
40	18.4	30.7	57	29	CNS020JW	AU206789 Tetraodon
41	18.4	30.7	69	28	BH408318	BH408318 1007054E
42	18.4	30.7	74	29	CC043506	CC043506 3591_1_15
43	18.4	30.7	77	10	BF303006	BF303006 602030933
44	18.4	30.7	78	9	AW075659	AW075659 xb29f08.x
45	18.4	30.7	79	12	RI083018	RI083018 602874538

#### ALIGNMENTS

RESULT 1  
BH413628  
LOCUS 1007033FC4.1EL\_x1 i007 - Rescuem Gr:d H Zea mays genomic, genomic  
DEFINITION survey sequence.  
ACCESSION BH413628  
VERSION BH413628.1 GI:17589985  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 88)  
AUTHORS Walbot.V.  
TITLE Maize genomic sequences found using engineered Rescuem transposon:  
JOURNAL Unpublished  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Very probable ligation site of ends cut by single endonuclease.  
Reverse complemented post-ligation sequence from source sequence.  
Plate: 1007033 column: 6

```

Class: transposon-tagged.
Location/Qualifiers
1..88
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1007 - RescueMu Grid H"
Note=Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site_1: BamHI; Site_2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid H was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."
BASE COUNT      17 a   29 c   32 g   :0 t
ORIGIN
Query Match          37.0%; Score 22.2; DB 28; Length 88;
Best Local Similarity 64.7%; Pred. No. 5.5e+03;
Matches 33; Conservative 0; Mismatches 18; Indels 0; Gaps 0

QY 1 GACATCCGCGAGGGGCCCAAGAGCCCTTCGCGACTACGTGGAGCGGTTTC 51
||||| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 38 GACAATCTTCCGCGAGGGGAGGCCCTTCGCGACTACGTGGAGCGGATC 88

RESULT 2
BH413629
LOCUS              88 bp    DNA        linear    GSS 12-DEC-2001
DEFINITION         Human Homo sapiens genomic clone ET33.5, genomic survey sequence.
ACCESSION           BH413629.1 GI:17589969
VERSION             BH413629.1
KEYWORDS            GSS.
SOURCE              Zea mays
ORGANISM            Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 88)
Waibot.V.
Maize genomic sequences found using engineered RescueMu transposon Unpublished
Contact: Waibot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: waibot@stanford.edu
Very probable ligation site found so sequence was trimmed.
Post-ligation sequence submitted separately.
Plate: 1007033 column: 6
Class: transposon-tagged.
Location/Qualifiers
1..88
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1007 - RescueMu Grid H"
Note=Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site_1: BamHI; Site_2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid H was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."
BASE COUNT      17 a   29 c   32 g   :0 t
ORIGIN
Query Match          37.0%; Score 22.2; DB 28; Length 88;
Best Local Similarity 64.7%; Pred. No. 5.5e+03;
Matches 33; Conservative 0; Mismatches 18; Indels 0; Gaps 0

QY 1 GACATCCGCGAGGGGCCCAAGAGCCCTTCGCGACTACGTGGAGCGGTTTC 51
||||| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 38 GACAATCTTCCGCGAGGGGAGGCCCTTCGCGACTACGTGGAGCGGATC 88

RESULT 3
AF039817/C
LOCUS              98 bp    DNA        linear    GSS 08-NOV-2000
DEFINITION         Human Homo sapiens genomic clone ET33.5, genomic survey sequence.
ACCESSION           AF039817.1 GI:3253280
VERSION             AF039817.1
KEYWORDS            GSS.
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 98)
Whitmore,S.A., Crawford,J., Apostolou,S., Eyre,H., Baker,E., Lower,K.M., Settassian,C., Goldup,S., Seshadri,R., Gibson,R.A., Mathew,C.G., Cleton-Jansen,A.M., Savola,A., Pronk,J.C., Auerbach,A.D., Doggett,N.A., Sutherland,G.R. and Callen,D.F.
Construction of a high-resolution physical and transcription map of chromosome 16q24.3: A region of frequent loss of heterozygosity in sporadic breast cancer
Genomics 50 (1), 1-8 (1998)
98292488
3628816
Contact: Whitmore SA
Department of Cyto Genetics and Molecular Genetics
Women's and Children's Hospital
72 King William Rd, North Adelaide, SA 5006, Australia
Email: whitmore@pulse.mad.adelaide.edu.au
Class: unknown
Location/Qualifiers
1..98
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="16q24.3"
/clone="EF33.5"
/clone_lib="Human"
BASE COUNT       31 a   21 c   34 g   12 t
ORIGIN
Query Match          36.3%; Score 21.8; DB 28; Length 98;
Best Local Similarity 70.7%; Pred. No. 7.4e+03;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0

QY 15 CCCAAGAGGCCTTCGCGAGCTACGTGGAGCGGCTTCITCA 55
||||| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 77 CGCTTGAGTCTCTTCGCCCTTCTTCGCCAGCTTCTCCA 37

```

```

RESULT 4
AC108059
LOCUS
DEFINITION
  AUI08059 Sugaco Homo sapiens cDNA library Homo sapiens cDNA clone
  2RV6C598, mRNA sequence.
ACCESSION
  AUI08059
VERSION
  AUI08059.1 GI:13557581
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 50)
  Suzuki.Y., Talra.H., Tsunoda.T., Mizushima-Sugano.J., Sese,J., Hata
  H., Ota.T., Isogai.T., Tanaka.T., Morishita,S., Okubo,K., Sakaki
  Y., Nakamura.Y., Suyama,A. and Sugano,S.
  Diverse transcriptional initiation revealed by fine, large-scale
  mapping of mRNA start sites
  EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE
  11375929
PUBMED
  11375929
COMMENT
  Contact: Yutaka Suzuki
  Department of Virology
  Institute of Medical Science, University of Tokyo
  4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
  Email: yusuk@ims.u-tokyo.ac.jp
  Suzuki.Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
  S. Construction and characterization of a full length-enriched and
  a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
  Location/Qualifiers
    1..50
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="2RV6C598"
    /clone_lib="Sugaco Homo sapiens cDNA library"
  BASE COUNT
    3 a 22 c 15 g 10 t
  ORIGIN
    Query Match 35.7%; Score 21.4; DB 9; Length: 50;
    Best Local Similarity 71.8%; Pred. No. 8.8e+03;
    Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

  QY 1 GACATCCGCGAGGCGCCAAAGAGGCGCTCCGCGACTAC 39
    1 ||||| 1 ||| 1 ||| 1 ||| 1 ||| 1 |||
  Db 12 GTCTCCGCGCGCGCGCGAGGCGCTGTCGCTCGAC 50
    1 ||||| 1 ||| 1 ||| 1 ||| 1 ||| 1 |||

RESULT 5
H27186
LOCUS
DEFINITION
  Y115d09.r1 Soares mouse p3NM1:9.5 Mus musculus cDNA clone
  IMAGE:350196 5', mRNA sequence.
  KD PROTEIN, HEAVY CHAIN (HUMAN);, mRNA sequence.
ACCESSION
  H27186
VERSION
  H27186.1 GI:897176
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 84)
  Hillier.L., Clark.N., Dubuque.T., Elliston.K., Hawkins.M., Holm
  M., Hultman.M., Kucaba.T., Le.M., Lennon.G., Marra.M., Parsons.J.,
  Rifkin.L., Rohlfing.T., Soares.M., Tan.F., Trevaskis.E., Waterston
  R., Williamson.A., Wohlmann.P. and Wilson.R.
  The WashU-Merck EST Project
  Unpublished
  Contact: Wilson RK
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63106
  Tel: 314 286 1800
  Fax: 314 286 1810

  TITLE
  JOURNAL
  COMMENT
  Email: mouseest@watson.wustl.edu
  This clone is available royalty-free through LNL : contact the
  IMAGE Consortium (infoimage.llnl.gov) for further information.
  MGI:221996
  Seq primer: mob.REGA+ET

  BASE COUNT
    17 a 31 c 15 g 16 t 5 others
  ORIGIN
    Query Match 35.7%; Score 21.4; DB 14; Length 84;
    Best Local Similarity 63.3%; Pred. No. 9.5e+03;
    Matches 31; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

  QY 6 CCGCAGGCGCCAAAGAGGCGCTCCGCGACTACGTGAGCGCTTCTTC 54
    1 ||||| 1 ||| 1 ||| 1 ||| 1 ||| 1 ||| 1 ||| 1 ||| 1 |||
  Db 30 CAGCCAGNGCTACGAGTCCAGCACCAAGATAGCTGACGGCTCTGC 78
    1 ||||| 1 ||| 1 ||| 1 ||| 1 ||| 1 ||| 1 ||| 1 ||| 1 |||

RESULT 6
W35927
LOCUS
DEFINITION
  W35927.1 GI:1317794
ACCESSION
  W35927
VERSION
  W35927.1
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 93)
  Marra.M., Hillier.L., Allen.M., Bowles.M., Dietrich.N., Dubuque.T.,
  Geisel.S., Kucaba.T., Lacy.M., Le.M., Martin.J., Morris.M.,
  Schellenberg.K., Steptoe.M., Tan.F., Underwood.K., Moore.B.,
  Theising.B., Wyllie.T., Lennon.G., Soares.B., Wilson.R. and
  Waterston.R.
  The WashU-HHMI Mouse EST Project
  Unpublished
  Contact: Marra M/Mouse EST Project
  WashU-HHMI Mouse EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: mouseest@watson.wustl.edu
  This clone is available royalty-free through LNL : contact the
  IMAGE Consortium (infoimage.llnl.gov) for further information.
  MGI:221996
  Seq primer: mob.REGA+ET

```

```

Email: est@watson.wustl.edu
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL : contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
Trace considered overall poor quality
Seq primer: M3Rp1
High quality sequence stop: 1.
Location/Qualifiers
  1..84
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="GDB:572239"
  /db_xref="taxon:9606"
  /clone="IMAGE:58321"
  /sex="female"
  /dev_stage="adult"
  /clone_lib="Soares breast 2NBHSt"
  /note="Organ: breast; Vector: pT73D (Pharmacia) with a
  modified polylinker; Site: 1; Not 1; Site 2: Eco RI; 1st
  strand cDNA was primed with a Not 1 - oligo(dT) primer [5'
  TGTACCAATCTGAAGTGGAGGCGCGCCCTTTTCTTTTCTTTT 3'],
  double-stranded cDNA was ligated to Eco RI adaptors
  (Pharmacia), digested with Not I and cloned into the Not I
  and Eco RI sites of a modified pT73 vector (Pharmacia).
  Library went through one round of normalization to a Cot =
  230. Library constructed by Bento Soares and M.Fatima
  Bonaído."
  BASE COUNT
    17 a 31 c 15 g 16 t 5 others
  ORIGIN
    Query Match 35.7%; Score 21.4; DB 14; Length 84;
    Best Local Similarity 63.3%; Pred. No. 9.5e+03;
    Matches 31; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

  QY 6 CCGCAGGCGCCAAAGAGGCGCTCCGCGACTACGTGAGCGCTTCTTC 54
    1 ||||| 1 ||| 1 ||| 1 ||| 1 ||| 1 ||| 1 ||| 1 ||| 1 |||
  Db 30 CAGCCAGNGCTACGAGTCCAGCACCAAGATAGCTGACGGCTCTGC 78
    1 ||||| 1 ||| 1 ||| 1 ||| 1 ||| 1 ||| 1 ||| 1 ||| 1 |||

RESULT 6
W35927
LOCUS
DEFINITION
  W35927.1 GI:1317794
ACCESSION
  W35927
VERSION
  W35927.1
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 93)
  Marra.M., Hillier.L., Allen.M., Bowles.M., Dietrich.N., Dubuque.T.,
  Geisel.S., Kucaba.T., Lacy.M., Le.M., Martin.J., Morris.M.,
  Schellenberg.K., Steptoe.M., Tan.F., Underwood.K., Moore.B.,
  Theising.B., Wyllie.T., Lennon.G., Soares.B., Wilson.R. and
  Waterston.R.
  The WashU-HHMI Mouse EST Project
  Unpublished
  Contact: Marra M/Mouse EST Project
  WashU-HHMI Mouse EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: mouseest@watson.wustl.edu
  This clone is available royalty-free through LNL : contact the
  IMAGE Consortium (infoimage.llnl.gov) for further information.
  MGI:221996
  Seq primer: mob.REGA+ET

```

High quality sequence stop: 84.

# FEATURES

Location/Qualifiers

1..93

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="IMAGE:350196"

/dev\_stage="9.5 dpc total fetus"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares mouse p3NMF19.5"

/notes="vector: pT73D (Pharmacia) with a modified

polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5,

TGTTACCAATCTCAAGTGGAGCGCGCATTTTITTTTITTTT 3',

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified pT73 vector

(Pharmacia). Library went through one round of

normalization to a Cot "5. Library constructed by Bonte

Soares and M.Fatima Bonaldo. KNA was kindly provided by

Dr. Minoru KO (Wayne State University)."

19 a 28 c 28 g 18 t

# BASE COUNT

ORIGIN

Query Match: 35.7%; Score 21.4; DB 14; Length 93;

Best Local Similarity 71.8%; Pred. No. 9.7e+04;

Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 19 ANGAGCCCTCCGGGACTACGCGCGGCTTCTTCAAG 57

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 44 AAGCGGTACTTGTGACTACTGTGACCGCTCTTCCAG 82

# RESULT 7

AZ920060

LOCUS

DEFINITION

1006017H08.x1 1006 - RescueMu Grid G Zea mays genomic, genomic

survey sequence.

ACCESSION

AZ920060

VERSION

AZ920060.1

KEYWORDS

GSS.

SOURCE

Zea mays

ORGANISM

Zea mays

REFERENCE

Walbot,V.

AUTHORS

Walbot V

TITLE

Maize genomic sequences found using engineered RescueMu transposon

JOURNAL

Unpublished

COMMENT

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 1006017 row: 37

Class: transposon-tagged.

Location/Qualifiers

1..92

/organism="Zea mays"

/mol\_type="genomic DNA"

/cultivar="mixed background W23/A189/B73"

/db\_xref="taxon:4577"

/tissue\_type="leaf"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="1006 - RescueMu Grid G"

/note="Organ: leaf; Vector: RescueMu (engineered from

pBlueScript backbone); Site.1: BamHI; Site.2: BglII;

RescueMu is a 4.9 kb, modified maize Mu transposon

designed to allow plasmid rescue from total genomic DNA.

Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdh.iastate.edu' and follow the links for 'RescueMu.' Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT 21 a 37 c 26 g 8 t

ORIGIN

Query Match: 34.7%; Score 20.8; DB 28; Length 92;

Best Local Similarity 64.6%; Pred. NO. 1.5e+04;

Matches 31; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GACATCCGCCGCGCCCAAGGAGCCCTCCCGGACTACGTGCGCCG 48

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 24 GACAACCTCTCCGCGACTACAAGAAGTCCGCGACTACGAGTCCCG 71

# RESULT 8

A1048494/c

LOCUS

DEFINITION

unh1a10.r1 Soares\_embryonic\_stem\_cell\_NMES Mus musculus cDNA clone

IMAGE:1749882 5' similar to SW:3BFL\_MOUSE P55194 SH3-BINDING

PROTEIN 3BP-1. ; mRNA sequence.

ACCESSION

A1048494

VERSION

A1048494.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

AUTHORS

(bases 1 to 58)

Marta,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisler,S., Kucaba,T., Lacy,M., Lee,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

The WashU-HMI Mouse EST Project

Unpublished

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:961694

Trace considered overall poor quality

Possible reversed clone; similarity of wrong strand

Seq primer: -28mi3 rev2 ET from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1..58

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="IMAGE:1749882"

/cell\_type="embryonic stem cell"

/lab\_host="DH10B"

/clone\_lib="Soares\_embryonic\_stem\_cell\_NMES"

/note="Organ: bone marrow; Vector: pT73D-pac (Pharmacia)

with a modified polylinker; Site.1: Not I; Site.2: Eco RI;

1st strand cDNA was primed with a Not I - oligo(dT) primer

[5

TGTTACCAATCTCAAGTGGAGCGCGCATTTTITTTTITTTTITTTTITTTT

3']; double-stranded cDNA was ligated to Eco RI adapters

(Pharmacia), digested with Not I and cloned into the Not

I and Eco RI sites of the modified pT73 vector. Library

went through two rounds of normalization, and was

```

BASE COUNT      9 a      16 c      26 g      7 t
ORIGIN
constructed by Bento Soares and M.Fatima Bonaldo."
Query Match      34.0%; Score 20.4; DB 9; Length 55;
Best Local Similarity 61.1%; Pred. No. 1.9e+04;
Matches 33; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GACATCCGCGAGCCGCCAAGAGCCCTTCGCGACTAGTGGACCGCTTCTTC 54
   ||| ||||| | | |||| | | || | | | | | | | | | | | | |
Db 56 GACCTCCGCGCGAGCCGCCAAGAGCCCTTCGCGACTAGTGGACCGCTTCTTC 3

RESULT 9
BQ760009      95 bp      mRNA      linear      EST 26-JUL-2002
LOCUS
DEFINITION
Ebp107_SQ002_G05_R pistil, 12 DPA, no treatment, cv Optic, Ebp107
Hordeum vulgare subsp. vulgare cDNA clone Ebp107_SQ002_G05 5', mRNA
sequence.
ACCESSION
BQ760009.1 GI:21968481
VERSION
BQ760009
SOURCE
Hordeum vulgare subsp. vulgare
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae
; Triticeae; Hordeum.
1 (bases 1 to 95)
REFERENCE
Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardile,L.,
Ramsay,L., Machray,G., Marshall,D.F.M. and Wauugh,R.
Development of Barley Transcriptome Resources
Unpublished
Contact: Waugh R. Marshall DF
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: estuwr@sari.ac.uk.
Location/Qualifiers
1..95
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Optic"
/db_xref="taxon:112509"
/clone="EBP107_SQ002_G05"
/tissue_type="pistil"
/dev_stage="12 DPA"
/clone_lib="DH10B"
/lab_host="pistil, 12 DPA, no treatment, cv Optic,
Ebp107"
/note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from pistils dissected from developing grains (12
days post anthesis) in glasshouse grown barley plants.
Developed as part of the barley transcriptome resources of
BSRC/SERAD funded cereal IGF (Investigating Gene
Function) project."
BASE COUNT      19 a      36 c      26 g      14 t
ORIGIN
Query Match      34.0%; Score 20.4; DB 13; Length 95;
Best Local Similarity 61.1%; Pred. No. 1.9e+04;
Matches 33; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 CATCCGCCAGGCGCCCAAGAGCCCTTCGCGACTAGTGGACCGCTTCTTCAA 56
   | |||| | || | | |||| | | | | | | | | | | | | | | |
Db 15 CTTCCGCGCGGCTCATCGAGCCCTCACTCCAGGCCCTGGCGCGAAGTACAA 68

RESULT 10
A1272982      74 bp      mRNA      linear      EST 21-DEC-1998
LOCUS

```

```

DEFINITION
q763c07.x1 NCI_CGAP_Utl Homo sapiens cDNA clone IMAGE:1986252 3',
mRNA sequence.
ACCESSION
A1272982
VERSION
A1272982.1 GI:3895250
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 74)
REFERENCE
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
TITLE
Tumor Gene Index
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Greg Lennon, Ph.D.
Cloning Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1779 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 67.
Location/Qualifiers
1..74
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1986252"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI-CGAP_Utl"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014"
BASE COUNT      14 a      27 c      17 g      16 t
ORIGIN
Query Match      33.7%; Score 20.2; DB 9; Length 74;
Best Local Similarity 63.3%; Pred. No. 2.1e+04;
Matches 31; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 12 GGGCCCAAGGAGCCCTTCGCGACTAGTGGACCGCTTCTTCAAGACC 60
   |||| | | | | | | | | | | | | | | | | | | | |
Db 14 GGGCCGAGGAGACCGCGCGCTCTTCTTGAATGCTTCAICCGGACC 62

RESULT 11
A1272996      77 bp      mRNA      linear      EST 13-SEP-1999
LOCUS
DEFINITION
df41b06.v1 Morton Retal Cochlea Homo sapiens cDNA clone
IMAGE:2485954 5', mRNA sequence.
ACCESSION
A1272996
VERSION
A1272996.1 GI:5876126
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 77)
REFERENCE
Robertson,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R.
and Morton,C.C.
Isolation of novel and known genes from a human fetal cochlear cDNA
library using subtractive hybridization and differential screening
Genomics 23, 42-50 (1994)
JOURNAL
MEDLINE
55130111
PUBMED
7829101

```



Insert length: 552 Std Error: 0.00

Seq primer: M13RP1

High quality sequence stop: 1.

Location/Qualifiers

## FEATURES

source

1. .91

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:490779"

/db\_xref="taxon:9606"

/clone="IMAGE:69482"

/sex="male"

/lab\_host="SOLR cells (kanamycin resistant)"

/clone\_lib="Stratagene placenta (#937225)"

/note="Organ: Placenta; Vector: Bluescript SK-; Site\_1:

ECORI; Site\_2: XhoI; Cloned unidirectionally. Primer:

Oligo dt. Caucasian. Average insert size: 1.2 kb; Uni-ZAP

XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGG 3' -3'

adaptor sequence: 5' CTCGACTTTTITTTTTT 3'

21 a 15 c 27 g 20 t 8 others

## BASE COUNT

ORIGIN

Query Match 33.7%; Score 20.2; DB 14; Length 9;

Best Local Similarity 63.6%; Pred. No. 2.2e+04;

Matches 28; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Oy 16 CCCAAGGAGCCCTCCCGGACTAGTGGACCGCTTCTTCAAGAC 59

Db 88 CCCANGATNCCCACCACTTACTGTCTCTCTCTCAGGAC 45

## RESULT: 14

CC043452/c

LOCUS

DEFINITION 3591\_l158\_l.G07\_2EL\_Y\_1 3591 - RescueMu Grid P Zea mays genomic,

genomic survey sequence.

VERSION CC043452

KEYWORDS GSS.

SOURCE CC043452.1 GI:29456343

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCA

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 75);

Walbot.V.

Maize genomic sequences found using engineered RescueMu transposon

Unpublished

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Possible ligation site of ends cut by 2 different endonucleases.

Reverse complemented post-ligation sequence from source sequence.

Plate: 3591\_l158\_l row: 29

Class: transposon-tagged.

Location/Qualifiers

1. .75

/organism="Zea mays"

/mol\_type="genomic DNA"

/cultivar="mixed background W23/A188/B73/K55"

/db\_xref="taxon:4577"

/tissue\_type="leaf"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="3591 - RescueMu Grid P"

/note="Organ: leaf; Vector: RescueMu (engineered from

pBluescript backbone); Site\_1: BamHI; Site\_2: BglII;

RescueMu is a 4.9 kb, modified maize Mu transposon

designed to allow plasmid rescue from total genomic DNA.

Mu elements insert preferentially into transcription

units. For more information on RescueMu, go to the web

site 'www.zmdb.iastate.edu' and follow the links for

'RescueMu.' Grid P was grown at Molokai in 2002. DNA was

extracted from leaf strips, double digested using BamHI

and BglII, and ligated to form circular plasmids. DH10B

cells were transformed and then screened on LB plates with

ampicillin."

BASE COUNT 6 a 29 c 24 g 16 t

## ORIGIN

Query Match 33.3%; Score 20; DB 29; Length 75;

Best Local Similarity 58.3%; Pred. No. 2.5e+04;

Matches 35; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Oy 1 GACATCCGCGAGGCCGCCAGGAGCCCTTCGCGACTAGTGGACCGCTTCTTCAAGACC 60

Db 68 GGCAACGTGCGCGCGCGAGTGGCGGTGGTGCAGAACACGTCGCGCAACGCGCAACACC 9

## RESULT 15

AA772825

LOCUS

DEFINITION

ae74el2.sl Stratagene schizo brain S11 Homo sapiens cDNA clone

IMAGE:969934 3' similar to TR:Q99613 Q99613 TRANSLATION INITIATION

FACTOR EIF-3 P110 SUBUNIT. [1] ; mRNA sequence.

AA772825

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished

Contact: Wilson HK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LIXL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Trace considered overall poor quality

Possible reversed clone: similarity on wrong strand

Possible reversed clone: polyT not found

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1. .79

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:969934"

/sex="male"

/tissue\_type="schizophrenic brain S-11 frontal lobe"

/dev\_stage="34 years old"

/lab\_host="SOLR (kanamycin resistant)"

/clone\_lib="Stratagene schizo brain S11"

/note="Vector: Bluescript SK-; Site\_1: EORI; Library

constructed from S-11 frontal lobe, male, 34 years old,

50% caucasian, 50% Aleutian. Schizophrenic suicide.

Random primed into EORI site of ZAP II Vector. Mass

excised. Custom library. Avg insert length 1.4kb.

Material obtained by Johnston N., Torrey, E.F., Volken R.,

and the Stanley Neuropathology Consortium - Analysis of

RNAs from the Brains of Individuals with Psychiatric

Diseases (Unpublished) Stanley Neurovirology Laboratory,

Johns Hopkins School of Medicine, Baltimore MD."

source

FEATURES





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:10:43 : Search time 144.5 Seconds  
(without alignments)  
1021.419 Million cell updates/sec

Title: US-09-475-704A-1

Perfect score: 60

Sequence: 1 gacatcaagcagdgcccaaa.....tggaccgattcttcaaaacc 60

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1660708 seqs, 1229959015 residues

Total number of hits satisfying chosen parameters: 1177748

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, NA:\*

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	21	35.0	87	9	US-09-864-761-28254 Sequence 23254, A
C 2	20.4	34.0	60	9	US-09-007-093-7 Sequence 7, Applt
C 3	19	31.7	60	12	US-09-908-975-16956 Sequence 16956, A
C 4	18.8	31.3	65	12	US-09-908-975-24532 Sequence 24532, A
C 5	18.6	31.3	66	9	US-09-815-242-1315 Sequence 1315, Ap
C 6	18.6	31.0	60	12	US-09-908-975-9744 Sequence 9744, Ap
C 7	18.6	31.0	65	12	US-09-908-975-25099 Sequence 25099, A
C 8	18.6	31.0	65	12	US-09-908-975-30489 Sequence 30489, A
C 9	18.4	30.7	70	14	US-10-060-036-424 Sequence 424, App
C 10	18.2	30.3	60	12	US-09-908-975-22576 Sequence 22576, A
C 11	18	30.0	60	12	US-09-908-975-14069 Sequence 14069, A
C 12	18	30.0	80	9	US-09-864-761-18009 Sequence 18009, A
C 13	17.8	29.7	60	12	US-09-908-975-13021 Sequence 13021, A
C 14	17.8	29.7	60	12	US-09-908-975-13510 Sequence 13510, A
C 15	17.8	29.7	60	12	US-09-908-975-20072 Sequence 20072, A
C 16	17.8	29.7	63	9	US-09-811-672-4 Sequence 4, Applt

```
Sequence 4, Applt
Sequence 1045, Ap
Sequence 336, App
Sequence 45, Appl
Sequence 119, App
Sequence 10, Appl
Sequence 10, Appl
Sequence 8727, Ap
Sequence 23836, A
Sequence 24465, A
Sequence 332, App
Sequence 332, App
Sequence 13049, A
Sequence 13049, A
Sequence 18261, A
Sequence 27199, A
Sequence 27227, A
Sequence 27227, A
Sequence 27703, A
Sequence 28587, A
Sequence 17, Appl
Sequence 16, Appl
Sequence 83464, A
Sequence 40, Appl
Sequence 13613, A
Sequence 914, App
Sequence 18498, A
Sequence 28077, A
Sequence 87, Appl
Sequence 88, Appl
Sequence 16, App
```

#### ALIGNMENTS

#### RESULT 1

```
US-09-864-761-28254/c
: Sequence 28254, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David K.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL I
: FILE REFERENCE: Aomico-X-1
: CURRENT APPLICATION NUMBER: US 09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.5
: PRIOR FILING DATE: 2003-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCI/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCI/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28254
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004870.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: P75080, EVALUJE 5.60e+00
US-09-864-761-28254

Query Match 35.0%; Score 21; DB 9; Length 87;
Best Local Similarity 73.0%; Pred. No. 2.5e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 20 AGAGGCCCTCCGGACTAGTGGACGCTTCITCAA 58
|||||||  ||  |||||  |||||  |||||  |||||
Db 86 AGAGGCCCAACCAACACTACTGACCAACATGATCAA 50

RESULT 2
US-09-007-093-7
; Sequence 7, Application US/09C07093
; Patent No. US2002025315A1
; GENERAL INFORMATION:
; APPLICANT: Anand, Naveen N
; APPLICANT: Barber, Brian H
; APPLICANT: Cates, George A
; APPLICANT: Caterini, Judith E
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DELIVERY OF
; TITLE OF INVENTION: ANTIGENS TO SELECTED CELLS OF THE IMMUNE SYSTEM
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McHurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/99/007,093
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,576
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael J
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-765

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCI/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCI/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28254
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004870.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: P75080, EVALUJE 5.60e+00
US-09-864-761-28254

Query Match 35.0%; Score 21; DB 9; Length 87;
Best Local Similarity 73.0%; Pred. No. 2.5e+02;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 20 AGAGGCCCTCCGGACTAGTGGACGCTTCITCAA 58
|||||||  ||  |||||  |||||  |||||  |||||
Db 86 AGAGGCCCAACCAACACTACTGACCAACATGATCAA 50

RESULT 2
US-09-007-093-7
; Sequence 7, Application US/09C07093
; Patent No. US2002025315A1
; GENERAL INFORMATION:
; APPLICANT: Anand, Naveen N
; APPLICANT: Barber, Brian H
; APPLICANT: Cates, George A
; APPLICANT: Caterini, Judith E
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DELIVERY OF
; TITLE OF INVENTION: ANTIGENS TO SELECTED CELLS OF THE IMMUNE SYSTEM
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McHurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/99/007,093
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,576
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael J
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-765

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-007-093-7

Query Match 34.0%; Score 20.4; DB 9; Length 60;
Best Local Similarity 65.2%; Pred. No. 4e+02;
Matches 30; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 13 GCCCCCAAGGACCCCTTCGGGACTACGTGGACCGCTTCTTCAAGA 58
|||||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 13 GGTCTAAAGACCTTTTAGAGACTATGTGTATAGGTTTATAAGA 58

RESULT 3
US-09-908-975-16956
; Sequence 16956, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPL
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16956
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-16956

Query Match 31.7%; Score 19; DB 12; Length 60;
Best Local Similarity 65.1%; Pred. No. 1.3e+03;
Matches 28; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 4 ATCAAGCAGGCGCCCAAGGAGCGCTTCGGCGGAC1ACGTGGACC 46
|||||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 5 AACGTCGCGCGCTCCACGCGGCTCTCCGGGACCTCGTCAGCC 47

RESULT 4
US-09-908-975-24532
; Sequence 24532, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPL
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
```

```
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24532
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-24532

Query Match      31.3%; Score 18.8; DB 12; Length 65;
Best Local Similarity 59.3%; Pred. No. 1.5e+03;
Matches 32; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 3 CATCAAGCAGGCCCAAGGAGCGCTTCGGCGACTACGTGGACCGCTTCTTCAAGAC 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 11 CCTAGAGCTGTGCCAGCAGCACTCTTCGCCGACCCTCTTCAACAGCAACAA 64

RESULT 5
US-09-815-242-1315/c
; Sequence 1315, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA-311A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/291,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,576
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1315
; LENGTH: 66
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-1315

Query Match      31.3%; Score 18.8; DB 9; Length 66;
Best Local Similarity 59.3%; Pred. No. 1.5e+03;
Matches 32; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 6 CAAGCAGGCCCAAGGAGCGCTTCGGCGACTACGTGGACCGCTTCTTCAAGAC 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 CAACCCGCGTTCAAAACCGCCCTTCCTCCCGCGCTCGCGCTCTCGGGCAAGCC 9

RESULT 6
US-09-908-975-9744
; Sequence 9744, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPI
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/257,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9744
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-9744

Query Match      31.0%; Score 18.6; DB 12; Length 60;
Best Local Similarity 65.9%; Pred. No. 1.8e+03;
Matches 27; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 18 CAAGGAGCCCTTCGGCGACTACGTGGACCGCTTCTTCAAGA 58
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 CAAGGAGACAGTCGCGACGAGTTCATCTTCTACTCCAGA 47

RESULT 7
US-09-908-975-25099
; Sequence 25099, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPI
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/257,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25099
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-25099

Query Match      31.2%; Score 18.6; DB 12; Length 65;
Best Local Similarity 61.2%; Pred. No. 1.8e+03;
Matches 30; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 11 AGGCCCAAGGAGCGCTTCGGCGACTACGTGGACCGCTTCTTCAAGAC 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 AGACACCCATGGAGATGCTTACCACCGTCGCGACGCTCAACATGTC 55

RESULT 8
US-09-908-975-30489
; Sequence 30489, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
```

APPLICANT: WASSERMAN, Alon  
APPLICANT: MINTZ, Eli  
APPLICANT: MINTZ, Liat  
APPLICANT: FAIGLER, Simchon  
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE  
FILE REFERENCE: 36688-0005  
CURRENT APPLICATION NUMBER: US/09/908,975  
CURRENT FILING DATE: 2001-07-20  
PRIOR APPLICATION NUMBER: US 60/287,724  
PRIOR FILING DATE: 2001-05-02  
PRIOR APPLICATION NUMBER: US 60/221,607  
PRIOR FILING DATE: 2000-07-28  
NUMBER OF SEQ ID NOS: 32337  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 30489  
LENGTH: 65  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-908-975-30489

Query Match 31.0%; Score 18.6; DB 12; Length 65;  
Best Local Similarity 57.9%; Pred. No. 1.8e+03;  
Matches 33; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

CY 1 GACATCAAGCAGGCGCCCAAGAGGCCCTTCCGCGACTACGTGCGACCGCTTCTTCAAG 57  
DB 3 GACCTTCCCAATGCCACCTGGACACCTTCCACCTCTCTGTCGATCTGGGAG 59

## RESULT 9

US-10-060-036-424  
Sequence 424, Application US/10060036  
Publication No. US20030073144A1  
GENERAL INFORMATION:  
APPLICANT: Benson, Darin R.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Persing, David H.  
APPLICANT: Hepler, William T.  
APPLICANT: Jiang, Yugu  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER  
FILE REFERENCE: 210121.566  
CURRENT APPLICATION NUMBER: US/10/060,036  
CURRENT FILING DATE: 2002-01-30  
NUMBER OF SEQ ID NOS: 4560  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 424  
LENGTH: 70  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc-feature  
LOCATION: 41..67  
OTHER INFORMATION: A = A,T,C or G  
US-10-060-036-424

Query Match 30.7%; Score 18.4; DB 14; Length 70;  
Best Local Similarity 58.5%; Pred. No. 2.1e+03;  
Matches 31; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

CY 1 GACATCAAGCAGGCGCCCAAGAGGCCCTTCCGCGACTACGTGCGACCGCTTCTTCAAG 57  
DB 3 GACATGAGGAGCTTCCACATGAAGCTTTCGTCGCCCANAGGAATCTCTACAI 55

## RESULT 10

US-09-908-975-22576/C  
Sequence 22576, Application US/09908975  
Publication No. US20030165843A1  
GENERAL INFORMATION:  
APPLICANT: SHOSHAN, Avi

APPLICANT: WASSERMAN, Alon  
APPLICANT: MINTZ, Eli  
APPLICANT: MINTZ, Liat  
APPLICANT: FAIGLER, Simchon  
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE  
FILE REFERENCE: 36688-0005  
CURRENT APPLICATION NUMBER: US/09/908,975  
CURRENT FILING DATE: 2001-07-20  
PRIOR APPLICATION NUMBER: US 60/287,724  
PRIOR FILING DATE: 2001-05-02  
PRIOR APPLICATION NUMBER: US 60/221,607  
PRIOR FILING DATE: 2000-07-28  
NUMBER OF SEQ ID NOS: 32337  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 22576  
LENGTH: 60  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-908-975-22576

Query Match 30.3%; Score 18.2; DB 12; Length 60;  
Best Local Similarity 58.2%; Pred. No. 2.5e+03;  
Matches 32; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

CY 3 CATCAAGCAGGCGCCCAAGAGGCCCTTCCGCGACTACGTGCGACCGCTTCTTCAAG 57  
DB 57 CATCAGGCTGGTCACCATGTGCGCTTGGCCCTTGATGCTTGGCTTCTTCAAG 3

## RESULT 11

US-09-908-975-14069  
Sequence 14069, Application US/0908975  
Publication No. US20030165843A1  
GENERAL INFORMATION:  
APPLICANT: SHOSHAN, Avi  
APPLICANT: WASSERMAN, Alon  
APPLICANT: MINTZ, Eli  
APPLICANT: MINTZ, Liat  
APPLICANT: FAIGLER, Simchon  
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE  
TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME  
FILE REFERENCE: 36688-0005  
CURRENT APPLICATION NUMBER: US/09/908,975  
CURRENT FILING DATE: 2001-07-20  
PRIOR APPLICATION NUMBER: US 60/287,724  
PRIOR FILING DATE: 2001-05-02  
PRIOR APPLICATION NUMBER: US 60/221,607  
PRIOR FILING DATE: 2000-07-28  
NUMBER OF SEQ ID NOS: 32337  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 14069  
LENGTH: 60  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-908-975-14069

Query Match 30.0%; Score 18; DB 12; Length 60;  
Best Local Similarity 60.0%; Pred. No. 3e+03;  
Matches 30; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

CY 8 ACACAGGCGCCCAAGAGGCCCTTCCGCGACTACGTGCGACCGCTTCTTCAAG 57  
DB 6 AGAAGGGTGGCCCTTCCGCGCGCTCCCAAGACCGCTGATGAACCTTCTTCAAG 55

## RESULT 12

US-09-864-761-18009/C  
Sequence 18009, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Pern, Sharon G.  
APPLICANT: Rank, David R.



```

RESULT 15
US-09-908-975-20072
: Sequence 20072, Application US/09908975
: Publication No. US20030155843A1
: GENERAL INFORMATION:
: APPLICANT: SHOSHAN, Avi
: APPLICANT: WASSERMAN, Alon
: APPLICANT: MINTZ, Eli
: APPLICANT: MINTZ, Zvi
: APPLICANT: FAIGLER, Simcha
: TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING
: FILE REFERENCE: 36686-0005
: CURRENT APPLICATION NUMBER: US/09/908,975
: PRIOR FILING DATE: 2001-07-20
: PRIOR APPLICATION NUMBER: US 60/267,724
: PRIOR FILING DATE: 2001-05-02
: PRIOR APPLICATION NUMBER: US 60/221,607
: PRIOR FILING DATE: 2000-07-28
: NUMBER OF SEQ ID NOS: 32337
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 20072
: LENGTH: 60
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-908-975-20072

```

```

Query Match      29.7%; Score 17.8; DB 12; Length 60;
Best Local Similarity 62.2%; Pred. No. 3.5e-03;
Matches 23; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

```

QY      15  CCCCAAGGAGCCCTCCGCGACTAGTGGAGCGGCTTCTTCAAGAC 59
      || ||||| | || | || | || |||| | ||||
Db       12  CCATCAGGAGTTCACGCGCAGCAACTTCCACCTCTTCATCCAGAC 56

```

```

Search completed: September 22, 2003, 17:30:57
Job time : 147.5 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic acid - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 : Search time 44 seconds  
(without alignments)  
601.886 Million cell updates/sec

Title: US-09-475-704a-1

Perfect score: 60

Sequence: 1 gacatcaagcagagccccc...tgacagcttc:caagacc 60

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 559978 sqqs, 220691566 residues

Total number of hits satisfying chosen parameters: 830498

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*

2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*

3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*

4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*

5: /cgn2\_6/ptodata/2/ina/PCRUS\_COMB.seq.\*

6: /cgn2\_6/ptodata/2/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	21.6	36.0	72	1 US-08-303-275-192
2	21.6	36.0	72	1 US-08-303-275-193
3	20.4	34.0	60	3 US-08-943-136-7
4	19.6	32.7	57	6 US-08-973-518-7
5	18.3	30.0	100	1 US-08-009-265-33
6	18.3	30.0	100	2 US-08-322-579-6
7	18.3	30.0	100	4 US-08-686-993A-18
8	17.8	29.7	40	1 US-08-495-243-25
9	17.8	29.7	40	1 US-08-548-078-3
10	17.8	29.7	40	1 US-08-495-739-25
11	17.8	29.7	40	1 US-08-495-741-25
12	17.8	29.7	40	3 US-08-062-023-25
13	17.8	29.7	63	3 US-08-750-419A-4
14	17.8	29.7	63	4 US-08-811-672-4
15	17.8	29.7	63	3 US-08-191-852-16
16	17.8	29.7	65	4 US-08-317-906-16
17	17.8	29.7	65	5 PCT-US95-13376-16
18	17.8	29.7	72	1 US-08-303-275-195
19	17.8	29.7	100	3 US-08-242-590A-45
20	17.6	29.3	50	4 US-08-554-329-119
21	17.6	29.3	72	1 US-07-778-233B-72
22	17.6	29.3	72	1 US-07-963-321-72
23	17.6	29.3	72	1 US-08-230-641-72
24	17.6	29.3	72	1 US-08-548-540-72
25	17.6	29.3	72	5 PCT-US96-09809-72
26	17.6	29.3	72	5 PCT-US96-2018-13
27	17.4	29.0	53	2 US-07-695-2018-13

C	28	17.4	29.0	53	3 US-08-470-532-13	Sequence 13, Appl
	29	17.2	28.7	50	1 US-08-575-052-7	Sequence 7, Appl
	30	17.2	28.7	50	1 US-08-614-516A-7	Sequence 7, Appl
	31	17.2	28.7	50	1 US-08-770-557-7	Sequence 7, Appl
	32	17.2	28.7	50	1 US-08-405-234-7	Sequence 7, Appl
	33	17.2	28.7	53	4 US-09-455-673-3	Sequence 3, Appl
	34	17.2	28.3	70	1 US-08-105-483-406	Sequence 406, App
	35	17.2	28.3	70	1 US-08-224-657-89	Sequence 89, Appl
	36	17.2	28.3	70	1 US-08-224-657-90	Sequence 90, Appl
	37	17.2	28.3	70	1 US-08-709-209-405	Sequence 406, App
	38	17.2	28.3	70	1 US-08-257-073-44	Sequence 44, Appl
	39	17.2	28.3	70	1 US-08-257-073-45	Sequence 45, Appl
	40	17.2	28.3	70	1 US-08-303-275-112	Sequence 112, App
	41	17.2	28.3	70	1 US-08-458-101-406	Sequence 406, App
	42	17.2	28.3	70	2 US-08-658-665-129	Sequence 129, App
	43	17.2	28.3	70	2 US-08-658-665-130	Sequence 130, App
	44	17.2	28.3	70	3 US-08-675-566-71	Sequence 71, Appl
	45	17.2	28.3	70	3 US-08-675-566-72	Sequence 72, Appl

## ALIGNMENTS

RESULT 1

US-08-303-275-192

Sequence 192, Application US/08303275

Patent No. 5766596

GENERAL INFORMATION:

APPLICANT: Paoletti, Enzo

APPLICANT: Tartaglia, James

APPLICANT: Cox, William I.

TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS RECOMBINANT

TITLE OF INVENTION: POXVIRUS VACCINE

NUMBER OF SEQUENCES: 205

CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtis, Morris & Saiford

STREET: 530 Fifth Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/303,275

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/897,382

FILING DATE: 11-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Frommer, William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454310-2420

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 192:

SEQUENCE CHARACTERISTICS:

LENGTH: 72 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-303-275-192

Query Match 36.0% Score 21.6; DB 1; Length 72;

Best Local Similarity 63.5% Pred. No. 1.4e+02;

Matches 33; Conservative 0; Mismatches 19; Indels 0; Gaps 0;





```

: ATTORNEY/AGENT INFORMATION:
: NAME: Stewart, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 2038-781
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 60 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-973-518-7

Query Match 34.0%; Score 20.4; DB 4; Length 60;
Best Local Similarity 65.2%; Pred. No. 3.2e+02;
Matches 30; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 13 GGCCCAAGAGAGCCCTCCGGGACTAGTGGACCGCTTCTTCAGAA 58
Db 13 GGTCTAAAGAACCTTTTAGAGACTATGTCGATAGGTTTATAAGA 58

RESULT 5
: Patent No. 5514566
: APPLICANT: FIDDES, JOHN C.; ABRAHAM, JUDITH A.
: TITLE OF INVENTION: METHODS OF PRODUCING RECOMBINANT
: FIBROBLAST GROWTH FACTORS
: CURRENT APPLICATION DATA:
: NUMBER OF SEQUENCES: 21
: FILING DATE: 05-APR-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/417,022
: FILING DATE: 16-DEC-1985
: APPLICATION NUMBER: 775,521
: FILING DATE: 12-SEP-1985
: SEQ ID NO: 17:
: LENGTH: 57
: 5514566-17

Query Match 32.7%; Score 19.6; DB 6; Length 57;
Best Local Similarity 62.0%; Pred. No. 5.8e+02;
Matches 31; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 AGGCCCCAAGAGCCCTTCGGGACTAGTGGACCGCTTCTTCAGAAC 60
Db 2 AGGACCCCAAGCCCTCTACTGCAAGAACGGGGCTTCTTCGAGAA 51

RESULT 6
: US-08-009-265-33
: Sequence 33, Application US/08009265
: Patent No. 5547871
: GENERAL INFORMATION:
: APPLICANT: Black Dr., Bruce C.
: APPLICANT: Summers Dr., Max D.
: TITLE OF INVENTION: Heterologous Signal Sequences For
: SECRETION: Secretion Of Insect Controlling Proteins
: NUMBER OF SEQUENCES: 53
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: American Cyanamid Company
: STREET: 1937 West Main Street, P.O. Box 60
: CITY: Stamford
: STATE: CT
: COUNTRY: USA
: ZIP: 06904-0060
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

: ATTORNEY/AGENT INFORMATION:
: NAME: Stewart, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 2038-781
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 60 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-973-518-7

Query Match 30.0%; Score 18; DB 1; Length 100;
Best Local Similarity 56.9%; Pred. No. 2e+03;
Matches 33; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 3 CATCAAGCAGGCGCCCTTCGGGACTAGTGGACCGCTTCTTCAGAAC 60
Db 31 CAACAATGANTGCACCAAGGTGCACCTGACAAAGGCTACTGTCCTTCTGCC 98

RESULT 7
: US-08-322-679-6
: Sequence 6, Application US/08322679
: Patent No. 5965123
: GENERAL INFORMATION:
: APPLICANT: Ahmed, Fakhruddin
: TITLE OF INVENTION: Coated Pesticidal Agents, Processes For
: PREPARATION: Their Preparation And Compositions Containing Them
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: American Cyanamid Company
: STREET: One Cyanamid Plaza
: CITY: Wayne
: STATE: New Jersey
: COUNTRY: United States
: ZIP: 07470
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/322,679
: FILING DATE: 13-OCT-1994
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Hogan, John W.
: REGISTRATION NUMBER: 32,703
: REFERENCE/DOCKET NUMBER: 32,458
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-831-3195
: TELEFAX: 201-831-3305
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 100 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-009-265-33
```



US-08-495-741-25

```

; LENGTH: 40 nucleotides
; TYPE: Nucleic acid
; STRANDEDNESS: single

```

```

; LENGTH: 40 nucleotides
; TYPE: Nucleic acid

```

LE  
TY



```

: TITLE OF INVENTION: RECOMBINANT ALLERGEN, FRAGMENTS THEREOF, CORRESPONDING RECOMBINAN
:
: TITLE OF INVENTION: MOLECULES, VECTORS AND HOSTS CONTAINING THE DNA MOLECULES, DIAGN
:
: TITLE OF INVENTION: THERAPEUTIC USES OF SAID ALLERGENS AND FRAGMENTS
:
: FILE REFERENCE: 1614-Q247P
:
: CURRENT APPLICATION NUMBER: US/03/811,572
:
: CURRENT FILING DATE: 2001-03-20
:
: NUMBER OF SEQ ID NOS: 28
:
: SOFTWARE: PatentIn version 3.1
:
: SEQ ID NO 4
:
: LENGTH: 63
:
: TYPE: DNA
:
: ORGANISM: Timothy Grass
:
: FEATURE:
:
: NAME/KEY: CDS
:
: LOCATION: (1)..(63)
:
: OTHER INFORMATION:
:
: US-03-811-672-4

```

```
Query Match:      29.7%   Score 17.8; DB 4; Length 63;
Best local Similarity: 58.5%; Pred. No. 2.2e+03;
Matches 31; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY  2 ACATCAAGCAGGGCCCGACGACGCCCTTCGCAGTACTACGTGGACGGCTTTTC 54
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  2 ACATCACCGACGACAGCGAGGAGGCCCATCGCCCTTACGACTTCGAGCTCTGC 54
```

Search completed: September 22, 2003, 14:10:33  
Job time : 47 secs

Genfore version 5.1.16  
Copyright (c) 1993 - 2003 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 14:02:05 : Search time 1798.5 Seconds  
(without alignment's)  
610.825 Million cell updates/sec

Title: us-09-475-704a-1

Perfect score: 60  
Sequence: 1 gacatcaagcagggccca.....tggacgttttcagacc 60

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2278192 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 452960

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match: 0%  
Maximum Match: 100%  
Listing first 45 summaries

Database : EST.\*

1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estlin.\*  
4: em\_estm.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hlc.\*  
9: gb\_estl.\*  
10: gb\_est2.\*  
11: gb\_hlc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_man.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_fod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*  
29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	21.8	36.3	98	28	AF039817
c 2	21.6	36.0	75	29	CC043452
c 3	21.4	35.7	93	14	W35927
c 4	20.6	34.3	88	28	BH413628

5	20.6	34.3	88	28	BH413629
6	20.2	33.7	74	9	A1272982
7	20.2	33.7	85	14	R12059
8	20.2	33.7	91	14	T48611
9	20	33.3	74	29	CC043506
10	20	33.3	79	9	AA772825
11	19.8	33.0	81	13	B0876237
12	19.8	33.0	84	14	H27186
13	19.6	32.7	86	9	AA119614
14	19.6	32.7	98	14	H45859
15	19.4	32.3	84	9	AA023110
16	19.2	32.0	70	28	AZ62487
17	19.2	32.0	84	28	BH222409
18	19.2	32.0	85	9	AA237114
19	19.2	32.0	88	9	AA909161
20	19.2	32.0	89	14	H44598
21	19.2	32.0	100	10	BF671124
22	19	31.7	70	9	A1036945
23	19	31.7	70	9	A1154612
24	19	31.7	82	9	AA615572
25	19	31.7	85	9	AA241182
26	18.8	31.3	50	9	AU102354
27	18.8	31.3	55	12	BM285363
28	18.8	31.3	55	12	BM433218
29	18.8	31.3	55	13	B0594077
30	18.8	31.3	64	9	A1026700
31	18.8	31.3	68	12	B1820594
32	18.6	31.0	73	14	H90720
33	18.6	31.0	74	29	CC249304
34	18.6	31.0	79	9	A1117457
35	18.6	31.0	86	9	AA627888
36	18.6	31.0	94	9	AA057469
37	18.6	31.0	99	9	AL829890
38	18.4	30.7	46	28	BH798543
39	18.4	30.7	67	29	CNS0203W
40	18.4	30.7	72	29	CC036825
41	18.4	30.7	83	9	AA117148
42	18.4	30.7	87	28	BH224249
43	18.2	30.3	47	28	BH807575
44	18.2	30.3	61	14	R71917
45	18.2	30.3	85	14	H43578

## ALIGNMENTS

RESULT :  
AF039817/c  
LOCUS AF039817 Human Homo sapiens genomic clone E133.5, genomic survey  
DEFINITION AF039817 Human Homo sapiens genomic clone E133.5, genomic survey  
sequence.  
ACCESSION AF039817  
VERSION AF039817.1 GI:3253280  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1 (bases 1 to 98)  
Whitmore, S.A., Crawford, J., Apostolou, S., Eyre-H., Baker, E., Lower, K.M., Sattasatian, C., Goldup, S., Seshadri, R., Gibson, R.A., Mathew, C.G., Cleton-Jansen, A.M., Savio, A., Pronk, J.C., Auerbach, A.D., Doggett, N.A., Sutherland, G.R. and Callen, D.F.  
Construction of a high-resolution physical and transcription map of chromosome 16q24.3: A region of frequent loss of heterozygosity in sporadic breast cancer.  
Genomics 50 (1), 1-8 (1998)

JOURNAL MEDLINE 98292488  
PUBMED 9628816  
COMMENT:  
Contact: Whitmore SA  
Department of Cytogenetics and Molecular Genetics  
Womens and Childrens Hospital  
72 King William Rd, North Adelaide, SA 5006, Australia



```

RESULT 4
BH413628
LOCUS
DEFINITION 1007033F04.y1 1007 - RescueMu Grid H Zea mays genomic, genomic
survey sequence.
ACCESSION BH413628
VERSION BH413628.1 GI:17589985
KEYWORDS GSS.
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 88)
AUTHORS Walbot,V.
JOURNAL
COMMENT
Maize genomic sequences found using engineered RescueMu transposon.
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site found so sequence was trimmed.
Post-ligation sequence submitted separately.
Plate: 1007033 column: 6
Class: transposon-tagged.
FEATURES
source
1..88
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1007 - RescueMu Grid H"
/notes="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site:1: BamHI; Site:2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu.' Grid H was grown at Berkeley in 2001. DNA
was extracted from leaf punches, double digested using
BamHI and BglII, and ligated to form circular plasmids.
DH10B cells were transformed and then screened on LB
plates with ampicillin."
BASE COUNT 17 a 29 c 32 g 10 t
ORIGIN
Query Match 34.3%; Score 20.6; DB 28; Length 88;
Best Local Similarity 62.7%; Pred. No. 1.2e+04;
Matches 32; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1 GACATCAAGCAGCGCCCAAGGAGCCCTTCGCGACTACGTGGACGCGCTC 51
||||| ||| ||| ||| ||| ||| ||| ||| |||
DB 38 GACATCTCGCGAGCGAGCGCCCTTCGCGACTACGTGGAGGATC 88
||||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
BH413629
LOCUS
DEFINITION 1007033F04.y1 1007 - RescueMu Grid H Zea mays genomic, genomic
survey sequence.
ACCESSION BH413629
VERSION BH413629.1 GI:17589989
KEYWORDS GSS.
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 88)
AUTHORS Walbot,V.
JOURNAL
COMMENT
Maize genomic sequences found using engineered RescueMu transposon.
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site found so sequence was trimmed.
Post-ligation sequence submitted separately.
Plate: 1007033 column: 6
Class: transposon-tagged.
FEATURES
source
1..88
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1007 - RescueMu Grid H"
/notes="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site:1: BamHI; Site:2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu.' Grid H was grown at Berkeley in 2001. DNA
was extracted from leaf punches, double digested using
BamHI and BglII, and ligated to form circular plasmids.
DH10B cells were transformed and then screened on LB
plates with ampicillin."
BASE COUNT 17 a 29 c 32 g 10 t
ORIGIN
Query Match 34.3%; Score 20.6; DB 28; Length 88;
Best Local Similarity 62.7%; Pred. No. 1.2e+04;
Matches 32; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1 GACATCAAGCAGCGCCCAAGGAGCCCTTCGCGACTACGTGGACGCGCTC 51
||||| ||| ||| ||| ||| ||| ||| ||| |||
DB 38 GACATCTCGCGAGCGAGCGCCCTTCGCGACTACGTGGAGGATC 88
||||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
BH413629
LOCUS
DEFINITION 1007033F04.y1 1007 - RescueMu Grid H Zea mays genomic, genomic
survey sequence.
ACCESSION BH413629
VERSION BH413629.1 GI:17589989
KEYWORDS GSS.
SOURCE
ORGANISM
Zea mays
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 74)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL
COMMENT
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.

```

```

ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 88)
AUTHORS Walbot,V.
JOURNAL
COMMENT
Maize genomic sequences found using engineered RescueMu transposon.
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site found so sequence was trimmed.
Post-ligation sequence submitted separately.
Plate: 1007033 column: 6
Class: transposon-tagged.
FEATURES
source
1..88
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1007 - RescueMu Grid H"
/notes="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site:1: BamHI; Site:2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu.' Grid H was grown at Berkeley in 2001. DNA
was extracted from leaf punches, double digested using
BamHI and BglII, and ligated to form circular plasmids.
DH10B cells were transformed and then screened on LB
plates with ampicillin."
BASE COUNT 17 a 28 c 33 g 10 t
ORIGIN
Query Match 34.3%; Score 20.6; DB 28; Length 88;
Best Local Similarity 62.7%; Pred. No. 1.2e+04;
Matches 32; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1 GACATCAAGCAGCGCCCAAGGAGCCCTTCGCGACTACGTGGACGCGCTC 51
||||| ||| ||| ||| ||| ||| ||| ||| |||
DB 38 GACATCTCGCGAGCGAGCGCCCTTCGCGACTACGTGGAGGATC 88
||||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
BH413629
LOCUS
DEFINITION 1007033F04.y1 1007 - RescueMu Grid H Zea mays genomic, genomic
survey sequence.
ACCESSION BH413629
VERSION BH413629.1 GI:17589989
KEYWORDS GSS.
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 74)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL
COMMENT
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.

```





BASE COUNT







PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV







XX (CHIR ) CHIRON CORP.  
 PA (UYST-) UNIV STELLENBOSCH.  
 XX Zur Megede J. Barnett SW, Etzge-brecht S, Van Rensburg RJ;  
 PI WPI; 2002-154920/20.  
 DR  
 XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful  
 PT in applications including DNA immunization or generation of packaging  
 PT cell lines, particularly in gene therapy.  
 XX Example 1; Fig 25; 233pp; Eng:ish.  
 PS  
 XX The present invention describes expression cassettes comprising a  
 CC polynucleotide sequence encoding a polypeptide comprising immunogenic  
 CC HIV type C polypeptides. The expression cassettes comprise any of the  
 CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpc, Env or  
 CC Nef (i). (i) have immunostimulant activity and can be used in gene  
 CC therapy. The HIV type C polynucleotides are useful in applications  
 CC including DNA immunisation, generation of packaging cell lines, and  
 CC production of HIV type C proteins. The polynucleotides are particularly  
 CC useful in gene therapy and DNA immunisation applications. AB03942 to  
 CC AB40054 and AB06204 to AB06215 represent sequences used in the  
 CC exemplification of the present invention.  
 XX  
 SQ Sequence 60 BP; 24 A; 13 C; 11 G; 12 T; 0 other;

Query Match 60.08; Score 36; DB 24; Length 60;  
 Best Local Similarity 75.08; Pred. No. 0.024;  
 Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
 QY 1 GACATCAAGCAGGCCCAAGGAGCCCTCCCGGACTAGTGAGCGCTTCTTCAAGACC 60  
 ||||| || || || || || || || || || || || || || || || || || || || || ||  
 Db 1 GACATAAAGCAAGGCCCAAGAGACCCCTTAGAGACTATGTCAGCGGCTCTTAAAGAC 60

RESULT 7  
 ABK36257  
 TO ABK36257 standard; DNA: 90 BP.  
 XX  
 AC ABK36257;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE HIV DNA encoding GAG segment 20.  
 XX  
 KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia; ds.  
 XX  
 OS Human immunodeficiency virus type 1.  
 OS Synthetic.  
 XX  
 PN WO200190197-A1.  
 XX  
 PD 29-NOV-2001.  
 XX  
 PF 25-MAY-2001; 2001WO-AU00622.  
 XX  
 PR 26-MAY-2000; 2000AU-0007761.  
 XX  
 XX (AUSU ) UNIV AUSTRALIAN NAT.  
 XX  
 PA Thomson SA, Ramshaw IA;  
 PI  
 XX WPI; 2002-147575/19.  
 DR  
 DR P-PSDB; AAU84418.  
 XX  
 XX New synthetic polypeptides having several different segments of at  
 PT least one parent polypeptide linked together differently compared to  
 PT the linkage in the parent polypeptide, for inducing immune response

PT against a pathogen or cancer -  
 XX  
 PS Example 1; Fig 12; 364pp; English.  
 XX  
 CC The invention relates to a new synthetic polypeptide (I) comprising  
 CC several different segments of at least one parent polypeptide linked  
 CC together in a different relationship relative to their linkage in the  
 CC parent polypeptide to impede, abrogate or otherwise alter at least one  
 CC function associated with the parent polypeptide and for inducing an  
 CC immune response against a pathogen or cancer. Also included are a  
 CC synthetic polynucleotide encoding and a computer system for  
 CC designing the synthetic polypeptides. The synthetic polypeptides and  
 CC polynucleotides are referred to as a Savine. The synthetic polypeptide is  
 CC useful for modulating immune responses preferably directed against a  
 CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
 CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
 CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
 CC Compositions comprising the polypeptide may be used in the treatment or  
 CC prophylaxis against viral (such as infections caused by HIV (human  
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
 CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,  
 CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic  
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
 CC sequence encodes a peptide derived from a parent protein used to  
 CC construct a savine of the invention.  
 XX

SQ Sequence 90 BP; 30 A; 20 C; 22 G; 14 T; 4 other;  
 Query Match 56.78; Score 34; DB 24; Length 90;  
 Best Local Similarity 74.18; Pred. No. 0.11;  
 Matches 40; Conservative 2; Mismatches 12; Indels 0; Gaps 0;  
 QY 7 AAGCAGGGCCCCAAGGAGCCCTCCCGGACTAGTGAGCGCTTCTTCAAGACC 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1 ARACAGGGAGCCCAAGAGAGCCCTTCAGAGACTATGTGGTAGGTTTTCAAAACC 54

RESULT 8  
 ABK36256  
 TO ABK36256 standard; DNA: 90 BP.  
 XX  
 AC ABK36256;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE HIV DNA encoding GAG segment 19.  
 XX  
 KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia; ds.  
 XX  
 OS Human immunodeficiency virus type 1.  
 OS Synthetic.  
 XX  
 PN WO200190197-A1.  
 XX  
 PD 29-NOV-2001.  
 XX  
 PF 25-MAY-2001; 2001WO-AU00622.  
 XX  
 PR 26-MAY-2000; 2000AU-0007761.  
 XX  
 XX (AUSU ) UNIV AUSTRALIAN NAT.  
 XX  
 PA Thomson SA, Ramshaw IA;  
 PI  
 XX WPI; 2002-147575/19.  
 DR  
 DR P-PSDB; AAU84417.  
 XX  
 XX New synthetic polypeptides having several different segments of at



CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at [ftp.wipo.int/pub/published\\_pat\\_sequences](http://ftp.wipo.int/pub/published_pat_sequences).  
XX  
SQ Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;  
  
Query Match 35.0%; Score 21; DB 22; Length 87;  
Best Local Similarity 73.0%; Pred. No. 1.4e+03;  
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
Qy 20 AGGAGCCCTCCCGGACTAGTGGACCGCTTCTTCAA 56  
||||| ||| ||||| || ||||| |||||  
Db 86 AGGAGCCCATCCACACTACCTGACCATCATGATCAA 50  
||||| ||| ||||| || ||||| |||||  
  
RESULT 11  
AAK20151/c  
ID AAK20151 standard; DNA: 87 BP.  
XX AC AAK20151;  
XX DT 05-NOV-2001 (first entry)  
XX DE Human brain expressed single exon probe SEQ ID NO: 20142.  
XX DE Human brain expressed single exon probe SEQ ID NO: 20142.  
XX DE Human: brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer; ss.  
XX CS Homo sapiens.  
XX PN WO200157275-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US00667.  
XX PR 04-FEB-2000; 2000US-0180312.  
XX PR 26-MAY-2000; 2000US-0207456.  
XX PR 30-JUN-2000; 2000US-0608408.  
XX PR 03-AUG-2000; 2000US-0632366.  
XX PR 21-SEP-2000; 2000US-0234667.  
XX PR 27-SEP-2000; 2000US-0236359.  
XX PR 04-OCT-2000; 2000GB-0024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI: 2001-483446/52.  
XX PT Single exon nucleic acid probes for analyzing gene expression in human  
XX brains -  
XX PS Example 4; SEQ ID NO: 20142; 550pp + Sequence Listing; English.  
XX CC The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX brain. They can be used to measure gene expression in brain cell samples,  
XX which may enable the diagnosis and improved treatment of nervous system  
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
XX epilepsy and cancers. The present sequence is one of the probes of the  
XX invention.  
XX SQ Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;  
  
Query Match 35.0%; Score 21; DB 22; Length 87;  
Best Local Similarity 73.0%; Pred. No. 1.4e+03;  
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
Qy 20 AGGAGCCCTCCCGGACTAGTGGACCGCTTCTTCAA 56  
||||| ||| ||||| || ||||| |||||  
Db 86 AGGAGCCCATCCACACTACCTGACCATCATGATCAA 50  
||||| ||| ||||| || ||||| |||||  
  
RESULT 12  
AAK46211/c  
ID AAK46211 standard; DNA: 87 BP.  
XX AC AAK46211;  
XX DT 06-NOV-2001 (first entry)  
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 20768.  
XX DE Human: bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukemia; lymphoma; myeloma; ss.  
XX CS Homo sapiens.  
XX PN WO200157276-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US00668.  
XX PR 04-FEB-2000; 2000US-0180312.  
XX PR 26-MAY-2000; 2000US-0207456.  
XX PR 30-JUN-2000; 2000US-0608408.  
XX PR 03-AUG-2000; 2000US-0632366.  
XX PR 21-SEP-2000; 2000US-0234667.  
XX PR 27-SEP-2000; 2000US-0236359.  
XX PR 04-OCT-2000; 2000GB-0024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI: 2001-488900/53.  
XX PT Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human bone marrow -  
XX PS Example 4; SEQ ID NO: 20768; 658pp + Sequence Listing; English.  
XX CC The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX bone marrow. They can be used to measure gene expression in bone marrow  
XX samples, which may enable the improved diagnosis and treatment of cancers  
XX such as lymphoma, leukemia and myeloma. The present sequence is one of  
XX the probes of the invention.  
XX SQ Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;  
  
Query Match 35.0%; Score 21; DB 22; Length 87;  
Best Local Similarity 73.0%; Pred. No. 1.4e+03;  
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
Qy 20 AGGAGCCCTCCCGGACTAGTGGACCGCTTCTTCAA 56  
||||| ||| ||||| || ||||| |||||  
Db 86 AGGAGCCCATCCACACTACCTGACCATCATGATCAA 50  
||||| ||| ||||| || ||||| |||||  
  
RESULT 13  
AAI52117/c  
ID AAI52117 standard; DNA: 87 BP.  
XX AC AAI52117;  
XX DT 17-OCT-2001 (first entry)  
XX DE Probe #20803 used to measure gene expression in human placenta sample.  
XX DE Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder; ss.  
XX

```

OS Homo sapiens.
XX WO200157272-A2.
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00663.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000US-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR:
XX WPI: 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta.
XX
XX Claim 25; SEQ ID No 20893; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
XX The present sequence is one such probe. The probes are useful for
XX producing a microarray for predicting, measuring and displaying gene
XX expression in samples derived from human placenta. The probes are useful
XX for antenatal diagnosis of human genetic disorders.
XX
XX Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;
XX
XX Query Match 35.0%; Score 21; DB 22; Length 87;
XX Best Local Similarity 73.0%; Pred. No. 1.4e+03;
XX Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
XX QY 20 AGGAGCCCTTCGCGACTACGTGACCGCTTCATCA 56
XX 86 AGGAGCCCATCCACACTACCTGAACCATGATCAA 50
XX
XX RESULT 14
XX ABS45930/c
XX ID ABS45930 standard; DNA; 87 BP.
XX
XX AC ABS45930;
XX
XX DT 25-FEB-2003 (first entry)
XX
XX DE Human liver single exon probe, SEQ ID No 20920.
XX
XX KW Human; single exon nucleic acid probe; liver; cirrhosis;
XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX coronary heart disease; SS.
XX
XX OS Homo sapiens.
XX
XX PN WO200157273-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00564.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000US-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR:
XX WPI: 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta.
XX
XX Claim 25; SEQ ID No 20893; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
XX The present sequence is one such probe. The probes are useful for
XX producing a microarray for predicting, measuring and displaying gene
XX expression in samples derived from human placenta. The probes are useful
XX for antenatal diagnosis of human genetic disorders.
XX
XX Sequence 87 BP; 17 A; 14 C; 29 G; 27 T; 0 other;
XX
XX Query Match 35.0%; Score 21; DB 22; Length 87;
XX Best Local Similarity 73.0%; Pred. No. 1.4e+03;
XX Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
XX QY 20 AGGAGCCCTTCGCGACTACGTGACCGCTTCATCA 56
XX 86 AGGAGCCCATCCACACTACCTGAACCATGATCAA 50
XX
XX RESULT 15
XX ABS20522/c
XX ID ABS20522 standard; DNA; 87 BP.
XX
XX AC ABS20522;
XX
XX DT 15-AUG-2002 (first entry)
XX
XX DE Human genome-derived single exon probe ORF from lung SEQ ID No 20513.
XX
XX KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familia; idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease; open reading frame; ORF.
XX
XX OS Homo sapiens.
XX
XX PN WO200186003-A2.
XX
XX PD 15-NOV-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00665.
XX
XX 04-FEB-2000; 2000US-180312P.
XX 26-MAY-2000; 2000US-207456P.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-234687P.
XX 27-SEP-2000; 2000US-236359P.
XX 04-OCT-2000; 2000US-0024263.

```

(MOLE-) MOLECULAR DYNAMICS INC.  
Penn SG, Hanzel DK, Chen W, Rank DR;  
WPI: 2002-114183/15.  
Spatially-addressable set of single exon nucleic acid probes, used to  
measure gene expression in human lung samples -  
Claim 4: SEQ ID NO 20513: 634pp: English.  
The invention relates to a spatially-addressable set of single exon  
nucleic acid probes for measuring gene expression in a sample derived  
from human lung comprising single exon nucleic acid probes having one of  
12614 nucleic acid sequences mentioned in the specification, or their  
complements or the 12387 open reading frames derived from the 12614  
probes. Also included are a microarray comprising the novel set of  
probes; the novel set of probes which hybridize at high stringency to a  
nucleic acid expressed in the human lung; measuring gene expression in a  
sample derived from human lung, comprising (a) contacting the array with  
a collection of detectably labeled nucleic acids derived from human lung  
mRNA, and (b) measuring the label detectably bound to each probe of  
the array; identifying exons in a eukaryotic genome, comprising  
(a) algorithmically predicting at least one exon from genomic sequences  
of the eukaryote; and (b) detecting specific hybridisation of detectably  
labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
having a fragment identical to the predicted exon, the probe is included  
in the above mentioned microarray; assigning exons to a single gene,  
comprising (a) identifying exons from genomic sequence by the method  
above and (b) measuring the expression of each of the exons in several  
tissues and/or cell types using hybridisation to a single exon  
microarrays having a probe with the exon, where a common pattern of  
expression of the exons in the tissues and/or cell types indicates that  
the exons should be assigned to a single gene; a peptide comprising one  
of 12011 sequences, mentioned in the specification, or encoded by the  
probes/open reading frames (ORF). The probes are used for gene  
expression analysis, and for identifying exons in a gene, particularly  
using human lung derived mRNA and for the study of lung diseases  
such as asthma, lung cancer, chronic obstructive pulmonary disease  
(COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
pulmonary alveolar proteinosis, Kargener syndrome, fibrocystic  
pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
and hyaline membrane disease. The present sequence is a single exon  
probe open reading frame of the invention.  
Note: The sequence data for this patent did not form part  
of the printed specification, but was obtained in electronic  
format directly from WPI at  
ftp.wipo.int/pub/published\_pat\_sequences.

Search completed: September 22, 2003, 14:08:49  
Job time : 195 secs